The University Of Sheffield.

Access To Thesis.

- A bound copy of every thesis which is accepted as worthy for a higher degree, must be deposited in the University of Sheffield Library, where it will be made available for borrowing or consultation in accordance with University Regulations.
- All students registering from 2008–09 onwards are also required to submit an electronic copy of their final, approved thesis. Students who registered prior to 2008–09 may also submit electronically, but this is not required.

Author: Johnathan Cooper-Knoc Dept: Neuroscience

Thesis Title: Characterisation of genotype-phenotype relationships in ALS associated with hexanucleotide repeat expansion of C9orf72 Registration No: 120223543

For completion by all students:

Submit in print form only (for deposit in the University Library): ☐

Submit in print form and also upload to the *White Rose eTheses Online* server: In full ☑

Edited eThesis ☐

Please indicate if there are any embargo restrictions on this thesis. Please note that if no boxes are ticked, you will have consented to your thesis being made available without any restrictions.

Embargo details: (complete only if requesting an embargo to either your print and/or eThesis)	Embargo required?		Length of embargo (in years)
Print Thesis	Yes ☐	No ☑	_____
eThesis	Yes ☐	No ☑	_____

Supervisor: I, the supervisor, agree to the named thesis being made available under the conditions specified above.

Name: Pamela J Shaw Dept: Neuroscience

Signed: ..... Date: 06/03/2015

Student: I, the author, agree to the named thesis being made available under the conditions specified above.

I give permission to the University of Sheffield to reproduce the print thesis in whole or in part in order to supply single copies for the purpose of research or private study for a non-commercial purpose.

I confirm that this thesis is my own work, and where materials owned by a third party have been used copyright clearance has been obtained. I am aware of the University's *Guidance on the Use of Unfair Means* (www.sheffield.ac.uk/lets/design/unfair)

I confirm that all copies of the thesis submitted to the University (including electronic copies on CD/DVD) are identical in content.

Name: Johnathan Cooper-Knock Dept: Neuroscience

Signed: ..... Date: 06/03/2015

For completion by students also submitting an electronic thesis (eThesis):

I, the author, agree that the University of Sheffield's eThesis repository (currently WREO) will make my eThesis available over the internet via an entirely non-exclusive agreement and that, without changing content, WREO may convert my thesis to any medium or format for the purpose of future preservation and accessibility.

I, the author, agree that the metadata relating to the eThesis will normally appear on both the University's eThesis server and the British Library's EThOS service, even if the thesis is subject to an embargo. I agree that a copy of the eThesis may be supplied to the British Library.

I confirm that the upload is identical to the final, examined and awarded version of the thesis as submitted in print to the University for deposit in the Library (unless edited as indicated above).

Name: Johnathan Cooper-Knock Dept: Neuroscience

Signed: ..... Date: 06/03/2015

THIS SHEET MUST BE BOUND IN THE FRONT OF THE PRINTED THESIS BEFORE IT IS SUBMITTED

Characterisation of genotype-phenotype relationships in ALS associated with hexanucleotide repeat expansion of C9orf72

Johnathan Cooper-Knock

Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield

March 2015

Submitted for degree of Doctor of Philosophy

I: Abstract

The papers selected represent characterisation of amyotrophic lateral sclerosis (ALS) patients belonging to the *C9orf72* genetic variant. The general introduction describes ALS more broadly, both clinically and pathologically, including a summary of proposed molecular mechanisms of pathogenesis. It goes on to discuss the discovery that GGGGCC-repeat expansions of *C9orf72* represent the most common genetic variant of this disease.

The first group of papers in section 3 relate to clinical and pathological characterisation of the *C9orf72* genetic variant. This includes genetic screening of cohorts of patients suffering from ALS, frontotemporal dementia (FTD), multiple sclerosis and parkinsonism.

The second group of papers in section 4 describes the use of transcriptome analysis, biochemical techniques and immunohistochemistry to study pathogenic mechanisms in *C9orf72*-ALS. Characterisation of the interactions and behaviour of RNA foci derived from the *C9orf72* repeat expansion suggest that these foci sequester proteins important to mRNA splicing. Novel methodology was then used to describe an increase in the splicing error rate in lymphoblastoid cell lines derived from *C9orf72*-ALS patients, which correlates with disease severity.

The final group of papers in section 5 describes the study of genetic modifiers of the C9orf72-disease phenotype. This includes the development and use of a Southern blotting protocol to size the expansion. This led to some interesting suggestions: that patients with intermediate length, but reputedly pathogenic, expansions do not exhibit haploinsufficiency or typical C9orf72-neuropathology. Finally contribution

was made to a larger study of the interaction between TMEM106B genotype and C9orf72-disease.

II: Acknowledgements

Permission for publication in this thesis of all of the described papers was obtained from the publishers and from all co-authors. In addition the relevant permissions were obtained for reproduction of material from the following review in which the candidate was a joint first author:

Cooper-Knock J, Kirby J, Highley R, Shaw PJ. The Spectrum of C9orf72-mediated Neurodegeneration and Amyotrophic Lateral Sclerosis. Neurotherapeutics. 2015 Mar 3.

On a more informal note, I would like to thank my supervisors Dr Janine Kirby, Prof Magnus Rattray and Prof Pamela Shaw for their invaluable support over the past five years as I have started out on a journey in science. I would also like to thank everybody in the team at the Sheffield Institute for Translational Neuroscience (SITraN), particularly Adrian Higginbottom, Guillaume Hautberge, Paul Heath and Robin Highley. I feel like the first stage of the journey is coming to an end, I have had fun, I hope you all have too.

III: Table of Contents

4. C9orf72-disease modifiers

    4.1. Expansion length

        4.1.1 Development of Southern blotting protocol (**Paper 9**)

        4.1.2 Effect of expansion length on disease phenotype (**Paper 8, Paper 10**)

    4.2. TMEM106B (**Paper 11**)

    4.3 Other disease modifiers

5. Concluding discussion – implications for translational work
6. References
7. Full list of published work

1. **Amyotrophic lateral sclerosis (ALS)**

1.1 Definition of ALS

ALS is a neurodegenerative disease defined clinically by loss of upper and/or lower motor neurons leading to progressive paralysis. It is a disease of aging: the peak age of onset is between 50 and 70 years (del Aguila *et al.* 2003). Death results from respiratory failure, usually between three and five years after onset (Haverkamp *et al.* 1995). ALS affects 2-3 individuals per 100,000 (Traynor *et al.* 1999) which has led to the perception that it is relatively rare. However, the prevalence is affected by the short survival time; the lifetime risk is estimated at 1 in 400 (Johnston *et al.* 2006) which is comparable to multiple sclerosis.

1.2 Clinical-pathological phenotype of ALS

The ALS phenotype is notably variable; for example approximately 20% of patients survive longer than 5 years (Preux *et al.* 1996) and despite the peak age of onset, cases have been reported in individuals as young as their third decade or as old as their tenth decade. The disease typically starts in one area and spreads throughout the motor system (Cooper-Knock *et al.* 2013). Most often this involves progression of weakness from one limb or the bulbar muscles to contiguous areas of the central nervous system (CNS); more rarely disease appears to initiate in the respiratory muscles or in multiple areas simultaneously. The rate at which additional areas become involved is a useful measure of the rate of disease progression (Roche *et al.* 2012). A significant proportion of ALS involves disease outside of the motor system. Most prominently there is an association with frontotemporal dementia (FTD): in general there is no overt cognitive dysfunction at disease onset, but as the disease

progresses cognitive impairment can develop in up to 50%, with clinically defined FTD occurring in 13-14% of cases (Phukan *et al.* 2012, Montuschi *et al.* 2014).

Pathologically ALS is characterised by degeneration of upper motor neurons in the motor cortex and lower motor neurons in the brainstem and spinal cord. In a similar manner to other neurodegenerative diseases of aging, such as Alzheimer's disease (AD) and Parkinson's disease (PD), affected neurons often contain cytoplasmic inclusions. The hallmark of these inclusions in sporadic ALS is misfolded ubiquitinated TDP-43 (Neumann *et al.* 2006). Indeed, the level of TDP-43 positive pathology has been shown to correlate with neuronal loss (Brettschneider *et al.* 2013, Brettschneider *et al.* 2014).

1.3 Aetiology of ALS

The vast majority of ALS is sporadic (SALS) i.e. it is not associated with a clear family history of disease. However, approximately 5-10% of cases have some family history, most often in an autosomal dominant pattern. There is some debate as to the definition of familial disease although the most widely accepted definition is 'a patient with ALS with either a first or second degree relative also with ALS' (Byrne *et al.* 2012). Primarily due to the advent of next generation sequencing we have seen a proliferation in discoveries of genetic causes of ALS such that, in northern Europe, ~70% of familial ALS (FALS) is now attributable to a particular genetic variant.

1.3.1 Genetic causes of ALS

A number of genetic variants of ALS have been identified (**Table 1**); the four most common genetic variants will be reviewed below:

Table 1: Major genetic variants associated with amyotrophic lateral sclerosis (ALS) (source http://alsod.iop.kcl.ac.uk/home.aspx)

Gene Symbol	Locus	Chromosome	Gene Symbol	Locus	Chromosome
SOD1	ALS1	21q22.11	*CHMP2B*	ALS17	3p12.1
ALS2	ALS2	2q33.2	*PFN1*	ALS18	17p13.3
ALS3	ALS3	18q21	*ERBB4*	ALS19	2q33.3-q34
SETX	ALS4	9q34.13	*HNRNPA1*	ALS20	12q13.1
SPG11	ALS5	15q14	*MATR3*	ALS21	5q31.2
FUS	ALS6	16p11.2	*CHCD10*	ALS-FTD2	22q11.23
ALS7	ALS7	20p13	*C9orf72*	ALS-FTD1	9p21.2
VAPB	ALS8	20q13.33	*UNC13A*	ALS	19p13.12
ANG	ALS9	14q11.1	*DAO*	ALS	12q24
TARDBP	ALS10	1p36.22	*DCTN1*	ALS	2p13
FIG4	ALS11	6q21	*NEFH*	ALS	22q12.1-q13.1
OPTN	ALS12	10p13	*PRPH*	ALS	12q12
ATXN2	ALS13	12q23-q24.1	*SQSTM1*	ALS	5q35
VCP	ALS14	9p13	*TAF15*	ALS	17q11.1-q11.2
UBQLN2	ALS15	Xp11.21	*SPAST*	ALS	2p24
SIGMAR1	ALS16	9p13	*ELP3*	ALS	8p21.1

1.3.1.1 SOD1-ALS:

Mutations in *SOD1* were the first identified genetic cause of ALS (Rosen *et al.* 1993). Mutations of *SOD1* account for 1.8-42.9% of all FALS cases depending on the population under consideration (Battistini *et al.* 2010). *SOD1* is a ubiquitously expressed protein with a normal role in free radical scavenging. It is relatively abundant throughout the CNS, especially within motor neurons (Pardo *et al.* 1995). Pathogenic mutations have been found throughout the length of the gene (Andersen 2006) but despite this, toxicity is thought to result from a gain-of-function (Gurney *et al.* 1996).

Clinically, *SOD1*-ALS is variable although certain clinical characteristics have been ascribed to certain *SOD1* mutations (Andersen 2006). For example, the most common *SOD1* mutation, D90A, is the only *SOD1* mutation to show autosomal recessive inheritance (Parton *et al.* 2002). *SOD1*-ALS is generally not associated with FTD (Wicks *et al.* 2009) although this has been disputed by a single case report (Katz *et al.* 2012).

Pathologically *SOD1*-ALS is distinct from other forms of ALS. Patients with *SOD1*-ALS display *SOD1* positive neuronal cytoplasmic protein aggregates which are negative for TDP-43 (Mackenzie *et al.* 2007). This has led to the suggestion that *SOD1*-ALS is distinct from other forms of ALS, which might go some way to explaining why treatments developed in *SOD1*-ALS animal models have generally not translated to the clinic (Aggarwal *et al.* 2008).

1.3.1.2 TARDBP-ALS:

Mutations in *TARDBP*, the ubiquitously expressed gene encoding TDP-43, have been identified in approximately 5% of patients with FALS (Sreedharan *et al.* 2008).

TDP-43 is predominantly located in the nucleus; it has been implicated in transcription regulation and mRNA processing including RNA splicing. TDP-43 contains 2 RNA recognition motifs (RRM), a nuclear localisation and nuclear export signal, and a glycine rich C-terminal region encoded by exon 6 of *TARDBP*, which contains the vast majority of disease causing mutations. This functional domain is thought to facilitate protein-protein interactions and to be essential for the function of TDP-43 in alternative splicing regulation (Buratti *et al.* 2005). In both sporadic ALS and *TARDBP*-ALS TDP-43 undergoes cytoplasmic mislocalisation and aggregation. These changes are not yet clearly understood: both loss of nuclear function and gain of cytoplasmic toxicity have been proposed and it may be that both are important (Halliday *et al.* 2012). It has been reported that the C-terminal domain is intrinsically aggregation prone and that this is enhanced by disease causing mutations (Johnson *et al.* 2009).

Clinically *TARDBP*-ALS, and even particular mutations of *TARDBP*, are reflective of the full spectrum of sporadic disease (Kirby *et al.* 2010). This suggests that other genetic and/or environmental modifiers are important. Unlike *SOD1* mutations, *TARDBP* mutations have also been found in patients with pure FTD and FTD-ALS (Pesiridis *et al.* 2009). Pathologically the most important feature of *TARDBP*-ALS is that it recapitulates the TDP-43 positive neuronal cytoplasmic inclusions found in the majority of the more numerous sporadic disease.

1.3.1.3 FUS-ALS:

Like *TARDBP*-ALS, *FUS*-ALS accounts for approximately 5% of patients with FALS (Vance *et al.* 2009). Also like TDP-43, fused in sarcoma (*FUS*) is a ubiquitiously expressed, predominantly nuclear, protein associated with RNA processing. In fact, FUS has a very similar structure to TDP-43 with 2 RNA recognition motifs (RRM), a

nuclear localisation and nuclear export signal, and a glycine rich C-terminal region, but unlike *TARDBP*-ALS, disease causing mutations are located throughout the gene, although in some cases pathogenicity remains to be validated (Da Cruz *et al.* 2011, Deng *et al.* 2014). *FUS*-ALS is also associated with cytoplasmic mislocalisation of the mutated protein although it is perhaps less absolute than for TDP-43 (Mackenzie *et al.* 2010). Moreover FUS has a recognised aggregation prone domain and FUS positive cytoplasmic aggregates are a feature of *FUS*-ALS cases, which do not exhibit TDP-43 positive inclusions. Better understanding of upstream mechanisms in ALS is likely to come from identification of common features of both *TARDBP*-ALS and *FUS*-ALS. For example, although both proteins have different RNA binding partners, they both target pre-mRNAs essential for neuronal integrity (Lagier-Tourenne *et al.* 2012).

Clinically *FUS*-ALS is a predominantly lower motor neuron disorder with variable upper motor neuron involvement and only rarely are FUS mutations associated with FTD (Blair *et al.* 2010). However, *FUS* positive neuronal cytoplasmic inclusions are a recognised feature of a subset of FTD patients without FUS mutations (Neumann *et al.* 2009).

1.3.1.4 C9orf72-ALS:

Perhaps the most exciting development in recent times has been the discovery of the genetic defect responsible for linkage on chromosome 9p21: intronic GGGGCC repeat expansions of *C9orf72* (DeJesus-Hernandez *et al.* 2011, Renton *et al.* 2011), which account for approximately 43% of FALS in the UK (Cooper-Knock *et al.* 2012). The frequency of this genetic variant means that study of a homogenous disease group is now possible in ALS. This genetic variant is the main focus of this thesis and will be explored in more detail in the next section.

1.3.2 Environmental causes of ALS

Twin studies suggest that risk of SALS is determined by both genetics and the environment (Al-Chalabi *et al.* 2010). Dietary toxins can cause motor neuron damage for example ingestion of the toxins responsible for lathyrism and konzo found in chickling peas and unprocessed cassava respectively, can result in spastic paraparesis (Ludolph *et al.* 1987, Howlett *et al.* 1990, Tylleskar *et al.* 1992). Dietary factors have been proposed to explain geographical clusters of ALS, such as on Guam (Spencer *et al.* 1987). Cigarette smoking is thought to confer an increased risk of ALS (Sutedja *et al.* 2007). Epidemiological studies have suggested a number of other risk factors, but results have been inconsistent. Some studies have suggested that high levels of physical activity could be a risk factor for ALS; one study found a particularly high prevalence of the disease amongst Italian professional footballers (Chio *et al.* 2005). However numbers were small and other studies investigating exercise have not found an association (Longstreth *et al.* 1998, Veldink *et al.* 2005).

It should be noted that with the discovery of an increasing number of genetic variants of ALS, screening has revealed that an increasing proportion of apparently sporadic ALS patients carry pathogenic mutations of presumably variable penetrance. This is especially true of *C9ORF72* expansions which are present in approximately 7% of apparently sporadic ALS cases in the UK (Cooper-Knock *et al.* 2012). It is possible that the majority of apparently sporadic ALS is the result of the interaction of a number of genetic risk factors of variable frequency. Genetic causes of neurodegenerative disease, because of the late age of onset, are excluded from reproductive pressure and therefore it is feasible that a large number of risk genes have persisted in the population.

1.3.3 Pathogenesis of ALS

The molecular pathogenesis of ALS is not well understood. Disease heterogeneity is a significant obstacle and is likely to explain in part the sheer number of proposed mechanisms. Much of the recent progress in this area has been made by examination of the newly discovered genetic variants of ALS and it is likely that *C9ORF72*-ALS, given the higher frequency of this variant, will herald a new era of understanding of ALS.

1.3.3.1 RNA processing

The discovery that TDP-43, a RNA binding protein, is the major component of neuronal cytoplasmic inclusions in ALS (Neumann *et al.* 2006) led to hypotheses involving dysfunctional RNA processing in ALS. Subsequently, mutations in a several RNA processing proteins have been discovered to cause familial ALS in certain cases including *FUS* (Vance *et al.* 2009), TBP associated factor 15 (*TAF15*) (Couthouis *et al.* 2011, Ticozzi *et al.* 2011), heterogeneous nuclear ribonucleoprotein A1 (*hnRNPA1)* and heterogeneous nuclear ribonucleoprotein A2B1 (*hnRNPA2B1*) (Kim *et al.* 2013). Alternative splicing is more prevalent in the CNS compared to other organs. This suggests that neurons may have particular requirements for RNA processing, and thus are selectively vulnerable to aberrant RNA processing. As well as determining what is dysfunctional in ALS, the normal function of RNA processing in neurons remains to be characterised. Perhaps the massive complexity in the nervous system is underpinned specifically by the divergent variability achievable through alternative splicing.

1.3.3.2 Protein processing

Cytoplasmic ubiquitinated aggregated protein inclusions are a hallmark of ALS and the quantity of TDP-43 pathology in ALS has been correlated with neuronal loss

(Brettschneider *et al.* 2013). This suggests that protein processing is dysfunctional in ALS. Additionally mutations in several genes encoding components of the protein degradation machinery, including ubiquilin 2 (*UBQLN2)* (Deng *et al.* 2011), valosin containing protein (*VCP*) (Johnson *et al.* 2010), optineurin (*OPTN)* (Hortobagyi *et al.* 2011) and sequestosome 1/p62 (*SQSTM1*) (Rubino *et al.* 2012), have been discovered in familial and sporadic ALS. *UBQLN2* and *VCP* relate specifically to the ubiquitin-proteasome system where individual proteins are targeted for destruction by ubiquitination. *OPTN* and *SQSTM1* are involved in the other major pathway for protein degradation, autophagy, which involves the sequestration of long-lived protein aggregates within autophagosomes. Fusion of the autophagosome with a lysosome leads to destruction of the aggregate. Both pathways have been implicated in the clearance of TDP-43 aggregations (Brady *et al.* 2011). There is some crossover between these systems and indeed *OPTN*, *SQSTM1* and *UBQLN2* all encode autophagy receptors which bind aggregates via ubiquitin and target them to the autophagosome (Thomas *et al.* 2013). Moreover, despite its role in the ubiquitin-proteasome system, *VCP* mutations have been shown to impair the fusion of autophagosomes with lysosomes (Tresse *et al.* 2010). Transfection of mutant *VCP* into mouse cortical primary neurons, as a means of disrupting function of the normal protein degradation pathways, results in TDP-43 mislocalisation and toxicity (Ritson *et al.* 2010) thus recapitulating major features of ALS.

1.3.3.3 Cytoskeleton and axonal transport

Motor neurons, which are relatively selectively targeted in ALS, have very specific anatomical features; their long projecting axons make them the largest cells in the human body. Transport of essential materials along axons requires molecular motors which travel along the cytoskeleton. Mutation of a number of genes encoding

components of the cytoskeleton have been implicated in ALS including profilin 1 (*PFN1*) (Wu *et al.* 2012), dynactin 1 (*DCTN1*) (Puls *et al.* 2003) and neurofilament heavy polypeptide (*NFH*) (Figlewicz *et al.* 1994).

1.3.3.4 Oxidative Stress

There is evidence for increased oxidative damage in tissues, including the CNS, from patients with ALS (Ferrante *et al.* 1997) and in the mutant *SOD1* mouse model (Ferrante *et al.* 1997). A recent study of lysine acetylation of TDP-43 showed that this covalent modification was promoted by arsenate-induced oxidative stress (Cohen *et al.* 2015). Moreover, the authors showed that as a result of acetylation of lysine residues present in its RRMs, TDP-43 molecules were reduced in their capacity to bind RNA and showed a propensity to form cytoplasmic aggregates; thus oxidative stress may modify the wild type TDP-43 protein in a way that mimics the effect of familial mutations.

1.3.3.5 Non-neuronal cells and Inflammation

There is growing evidence for the involvement of non-neuronal cells in the propagation of ALS through the CNS, if not the initiation of disease. Neuroinflammation and activation of microglia has been detected pathologically (Kawamata *et al.* 1992) and in imaging studies (Turner *et al.* 2004) of ALS patients. However, it has proven difficult to determine whether the association is protective or harmful or both. Microglia, the resident macrophages of the CNS, can adopt either a neuroprotective (M2) or a neurotoxic (M1) phenotype depending on their surroundings which includes modulation by other inflammatory cells such as astrocytes and T-lymphocytes (Mantovani *et al.* 2004). Whilst some studies have suggested that microglial activation correlates with neuronal death (Turner *et al.*

2004), microglial activation has been shown to precede disease onset in mutant *SOD1* mice (Alexianu *et al.* 2001) and indeed it has been suggested that early in disease microglia might exist in the neuroprotective M2 phenotype. Crucially, microglia appear to become neuroprotective in response to CD4+ T-cell signalling (Appel *et al.* 2010). Crossing mutant *SOD1* mice with CD4 -/- knockout mice reduced survival (Beers *et al.* 2008) and was associated with reduction in the numbers of M2-type microglia. Alternatively addition of activated CD4+ T-cells from wild type donor mice to the mutant *SOD1* mice delayed motor neuron death and improved survival (Banerjee *et al.* 2008).

It has been suggested that non-neuronal cells may be responding to the presence of protein aggregates. Using recombinant mutated *SOD1* protein in primary cultures of motor neurons and microglia, extracellular m*SOD1* has been demonstrated to functionally activate microglia and, although m*SOD1* is not directly toxic to motor neurons alone, it is toxic in the presence of a co-culture with microglia (Zhao *et al.* 2010). This suggests that misfolded protein may be responsible for disease propagation via activation of non-neuronal cells. This may be amplified by prion-like induction of misfolding in normal protein, and if neurons affected by disease spread in this way produce additional aggregated protein, a positive-feedback loop would be established.

1.3.3.6 Mitochondrial Function

As well as energy metabolism mitochondria are essential for calcium homeostasis, control the initiation of apoptosis and are the major source of reactive oxygen species (Cozzolino et al. 2012). All of these mechanisms have been implicated in ALS. Moreover mitochondria are damaged in excitotoxicity, which is the target of riluzole, the only proven disease modifying agent in MND (Bensimon et al. 1994).

Mitochondria in ALS are morphologically and functionally abnormal in both patients and in mouse models (Cozzolino et al. 2012).

1.4 Conclusion

Clinically and pathologically ALS has been well characterised, although the recent discovery of multiple genetic variants of the disease has necessitated characterisation of the distinct features of each subtype. Meanwhile, understanding of the molecular pathogenesis of ALS has been progressing at a rapid rate, in large part as a result of increased understanding of the genetic basis of ALS. We have highlighted a number of pathophysiological mechanisms and explored possible links between them. Interestingly, the role of non-neuronal cells appears to interact with that of protein misfolding, perhaps via a positive feedback loop; if the production of misfolded protein cannot be stopped then modulation of the downstream inflammatory response may be successful.

2. **Clinical and pathological characterisation of *C9orf72*-disease**

2.1 Clinical presentation of C9orf72-ALS

Paper 1 describes the clinical characterisation of 62 *C9orf72*-ALS cases from a cohort of 563 cases in Northern England (Cooper-Knock *et al.* 2012), and **paper 2** describes a cross-sectional study bringing together data from various sites globally (Majounie *et al.* 2012). Whilst the full spectrum of the ALS clinical phenotype is represented within *C9orf72*-ALS, the most significant and robust clinical feature associated with patients carrying a repeat expansion is the increased incidence of FTD or a family history of FTD in up to 50% of cases (**paper 1**, (Byrne *et al.* 2012, Chio *et al.* 2012, Cooper-Knock *et al.* 2012, Sabatelli *et al.* 2012, Stewart *et al.* 2012, Garcia-Redondo *et al.* 2013)). This is not surprising given that the 9p21 risk locus was initially identified through mapping ALS-FTD families, where cases presented with either ALS or FTD or both diseases (Hosler *et al.* 2000, Mok *et al.* 2012). There is also an increase in the incidence of bulbar onset in *C9orf72*-ALS of up to 44%, compared to an average of 25-26% in non-*C9orf72* ALS (**paper 1**, (Chio *et al.* 2012, Cooper-Knock *et al.* 2012, Stewart *et al.* 2012)), and several groups also found evidence of an earlier age of onset by 1.8-5yrs (Byrne *et al.* 2012, Sabatelli *et al.* 2012, van Rheenen *et al.* 2012, Garcia-Redondo *et al.* 2013). *C9orf72*-ALS has also been associated with a shorter disease duration by 5.7-12 months suggesting a more aggressive disease course (**paper 1**, (Byrne *et al.* 2012, Cooper-Knock *et al.* 2012, Sabatelli *et al.* 2012, Garcia-Redondo *et al.* 2013)).

Whilst not all clinical cohorts show all of these characteristics, the inconsistencies may not only be due to the different populations under consideration, but also due to the groups under comparison, as some reports compare *C9orf72*-ALS with all non-*C9orf72* ALS cases, whereas others compare *C9orf72*-ALS specifically with familial

or sporadic non-*C9orf72* ALS. This is highlighted in a Belgian cohort of ALS patients, where comparing familial *C9orf72*-ALS patients with non-*C9orf72* FALS cases revealed phenotypic differences but comparing sporadic *C9orf72*-ALS cases to non-*C9orf72* sporadic ALS did not (Debray *et al.* 2013). Similarly, Millicamps and colleagues compared the clinical phenotype of *C9orf72*-ALS with that of known *SOD1*-ALS, *TARDBP*-ALS and *FUS*-ALS patients (Millecamps *et al.* 2012): bulbar onset was found more frequently in *C9orf72*-ALS than in the other three genetic variants but *C9orf72*-ALS cases had shorter disease duration compared only to *SOD1*-ALS and *TARDBP*-ALS cases, and they had an older age of onset compared only to *SOD1*-ALS and *FUS*-ALS cases.

Interestingly in **paper 1**, within our screened cohort of 361 normal controls, two individuals were found to have a *C9orf72* expansion of >30 repeats (Cooper-Knock *et al.* 2012). Neither case had any relevant past medical or family history. In **paper 2**, five healthy controls from 2585 screened were shown to carry the *C9orf72* expansion. Of course these cases may be presymptomatic but many of them were within or beyond the peak age of onset, and therefore it appears likely that the *C9orf72* expansion has variable penetrance. This is not surprising given that a proportion of *C9orf72*-ALS patients have no family history of disease. In **paper 1,** we explored the possibility that these apparently sporadic patients were actually undisclosed familial patients: Sixteen of the 35 sporadic C9orf72-ALS patients either had a family history of neurological disease which was not ALS, or at least one parent died at <70 years; but for the remaining 19 patients there was no suggestion of familial disease, consistent with less than 100% penetrance. Based on the frequencies reported in **paper 1** and **paper 2**, up to 0.5% of control individuals may carry the expansion. It is notable that none of controls shown to have an expansion

in **paper 2** were aged >80 years; thus it is possible that the expansion may become 100% penetrant if an affected individual lives sufficiently long.

In **paper 1** we describe two cases with potentially pathogenic mutations in addition to an expansion of *C9orf72:* one patient had a p.Ala321Val mutation in *TARDBP* and another had a p.Glu322Lys mutation in *OPTN.* Another study has shown that the number of *C9orf72*-disease individuals with a second described pathogenic mutation is higher than would be expected by chance (van Blitterswijk *et al.* 2013). This is an interesting observation which may have implications for penetrance – perhaps a 'second hit' is necessary to initiate disease.

In **paper 2**, it was shown that the *C9orf72* expansion was associated with at least some component of the 9p21 risk haplotype in all cases examined. This has led to the proposal that a common founder is responsible for all of *C9orf72*-disease (Majounie *et al.* 2012) but this is controversial (Dobson-Stone *et al.* 2012, Beck *et al.* 2013, Smith *et al.* 2013, Cooper-Knock *et al.* 2014). Observations of population frequencies in **paper 2** are consistent with a common founder in Scandinavia: in Finland *C9orf72* expansions were found in 61% of FALS patients (Majounie *et al.* 2012); but further away from Scandinavia the expansion frequency becomes less: we showed that *C9orf72* expansions are present in 43% of FALS cases in Northern England (**paper 1**, (Cooper-Knock *et al.* 2012)); in Germany 22% of FALS cases carry the expansion (**paper 2**, (Majounie *et al.* 2012)) but in Japan the equivalent Figure is only 3.4% (Konno *et al.* 2013). It is possible that it is not the expansion itself which is inherited but a propensity for the region to expand. One of the initial studies of *C9orf72* noted that the risk haplotype was associated with an increased number of repeats even in controls (Renton *et al.* 2011). More recently it has been shown that a 9p21 haplotype is significantly associated with sporadic ALS even if

patients with the GGGGCC-repeat expansion are excluded (Jones *et al.* 2013); perhaps an additional, as yet unidentified, repeat sequence is present at the same locus.

As well as FTD, a number of other neurological diseases were found in *C9orf72*-ALS probands or their pedigrees in our cohort (**paper 1**). 6.5% of patients either suffered or had a family history of Parkinson's disease (PD); 5% of patients either suffered co-morbid demyelination or had a family history of demyelinating disease. This formed the basis for our work in **paper 3** (see section *3.4*) and **paper 4** (see section *2.5*).

2.2 Neuropathology of C9orf72-ALS

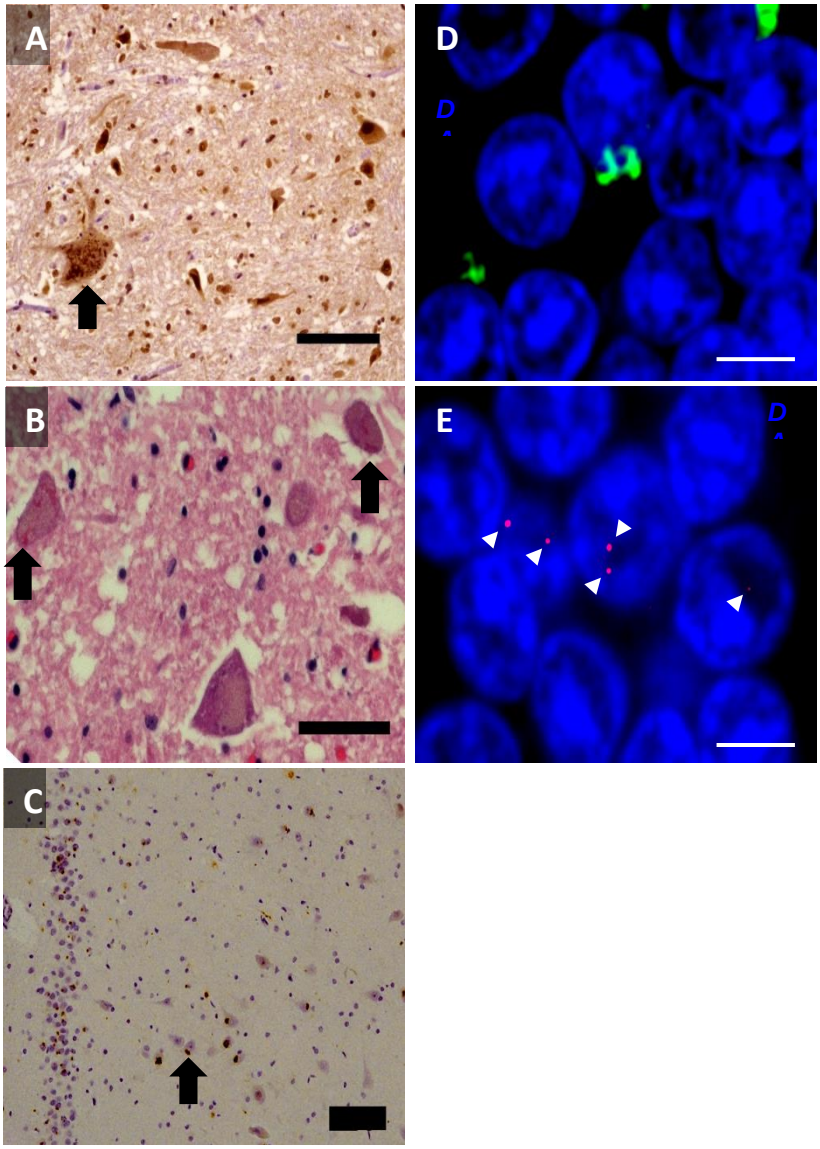
Pathological examination as part of **paper 1** revealed that the *C9orf72* expansion is associated with classical Bunina bodies, p62 and TDP-43 positive neuronal and glial cytoplasmic inclusions in the motor cortex and anterior horns of the spinal cord, and with marked loss of motor neurons (Cooper-Knock *et al.* 2012) (**Figure 1A-B**). Thus, *C9orf72*-ALS is a TDP-43 proteinopathy resembling sporadic ALS more broadly. However, the repeat expansion cases have additional characteristic pathology in extra-motor areas: p62 and ubiquitin positive but TDP-43 negative neuronal cytoplasmic inclusions (NCI) were identified by ourselves and others in the three layers of the cerebellar cortex, the hippocampus, and the neocortex (Murray *et al.* 2011, Cooper-Knock *et al.* 2012, Stewart *et al.* 2012) (**Figure 1C**). In addition, neuronal intra-nuclear inclusions (NII) were found in the cerebellar granular cells and hippocampal pyramidal cells (Al-Sarraj *et al.* 2011). The pathology in these extra-motor regions, which were initially not thought to be affected in ALS, were first described in FTD and FTD-MND cases (Pikkarainen *et al.* 2008, King *et al.* 2009) and subsequently in ALS (King *et al.* 2011), before being associated with the

C9orf72 expansion (Al-Sarraj *et al.* 2011). The NCIs and NIIs, as well as being positive for ubiquitin and p62, also stain for dipeptide repeat proteins (DPRs), which are thought to be generated through repeat associated non-ATG (RAN) translation of repeat-RNA transcribed from the C9orf72 expansion (Mann *et al.* 2013). DPRs derived via sense transcription of the expansion (poly-GA, poly-GR and poly-GP) generally co-localise with p62 positive NCI in the cerebellum and hippocampus (**Figure 1D**). DPRs are also derived via antisense transcription of the expansion (poly-PA, poly-PR and poly-GP) (Mori *et al.* 2013). Both sense and antisense DPRs have been identified in motor areas including spinal motor neurons (Mori *et al.* 2013). Finally, all tissues thus far examined in patients carrying *C9orf72* expansions contain nuclear RNA foci transcribed directly from the *C9orf72* expansion in both a sense and an antisense direction (see section on RNA toxicity) (Lagier-Tourenne *et al.* 2013, Mizielinska *et al.* 2013, Cooper-Knock *et al.* 2014) (**Figure 1E**). The neuropathology of C9orf72-ALS including the distribution of DPRs and RNA foci is developed further in **paper 5** and **paper 6** as described in section *4.*

2.3 Clinical presentation of C9orf72-FTD

In **paper 2,** screening of FTD cases found the GGGGCC-repeat expansion of *C9orf72* was found in 25.1% (range 13.8% to 100% depending on population) of familial FTD and 5.8% (range 0% to 18.8% depending on population) of sporadic FTD cases worldwide (Majounie *et al.* 2012). Whilst FTD can present as behavioural variant FTD (bvFTD), progressive non-fluent aphasia (PFNA) or semantic dementia (SD), those with the *C9orf72* expansion predominantly present with bvFTD including progressive personality deterioration, such that affected individuals may exhibit psychosis in the form of hallucinations and delusions (**paper 2**, (Majounie *et al.* 2012)). The frequency of bvFTD is consistently higher in *C9orf72* FTD patients than

Figure 1 (reproduced from Cooper-Knock, Kirby et al): Characteristic Pathology of *C9orf72*-ALS. Motor neurons of the spinal cord display typical TDP-43 pathology including cytoplasmic TDP-43 positive skeins and compact inclusions (A, anterior horn, DAB stained for pTDP-43, scale bar 100 µm) and Bunina bodies (B, anterior horn, H+E stain, scale bar 100 µm). However, in addition *C9orf72*-ALS patients display p62-positive cytoplasmic inclusions in extra-motor areas (C, Hippocampus dentate gyrus, DAB stained for p62, scale bar 100µm) which also stain for dipeptide repeat protein (D, cerebellar granule neurons, stained for poly-GA and DAPI, scale bar 3µm). Numerous tissues also show nuclear RNA foci (E, cerebellar granule neurons, stained for $(GGGGCC)_3$ and DAPI, foci are indicated by arrowheads, scale bar 3µm).



in non-*C9orf72* FTD cases in multiple populations across the world (Dobson-Stone *et al.* 2012, Sha *et al.* 2012, Snowden *et al.* 2012, Kaivorinne *et al.* 2013). Some patients with *C9orf72* expansions do present with PNFA with loss of word retrieval and non-fluent speech culminating in loss of speech, but the frequency of this presentation is similar in both *C9orf72* and non-*C9orf72* cohorts (Kaivorinne *et al.* 2013). In contrast, SD, where individuals lose their understanding of words and objects, is only rarely associated with *C9orf72* expansion (Simon-Sanchez *et al.* 2012, Snowden *et al.* 2012).

2.4 C9orf72 expansions and parkinsonism

We (**paper 1**) and others noted an apparent increase in the incidence of PD, parkinsonism concomitant with ALS or a family history of PD in screening of ALS cases for the *C9orf72* expansion (Boeve *et al.* 2012, Cooper-Knock *et al.* 2012). In **paper 3** we performed screening for the *C9orf72* expansion in a cohort of PD patients. We and others identified a few rare incidences of the repeat expansion, usually in cases with atypical PD (Cooper-Knock *et al.* 2013, Lesage *et al.* 2013, Lindquist *et al.* 2013). Further reports failed to find *C9orf72* expansions over 30 repeats in PD (Jiao *et al.* 2013, Nuytemans *et al.* 2013). However, both of these studies identified an increase in the number of PD patients with intermediate length expansions (defined as 7-30 repeats), compared to controls.

To further explore this issue, in **paper 3** we conducted a pathological examination of the substantia nigra in *C9orf72*-ALS cases. We identified an increased number of p62 positive and TDP-43 negative NCIs compared to non-*C9orf72* cases, and this was associated with a marked loss of dopaminergic neurons (Cooper-Knock *et al.* 2013). Thus we suggest that the increased incidence of parkinsonism in *C9orf72*-

disease appears to be due to a direct effect of the *C9orf72* pathological process on the substantia nigra neurons rather than classical alpha-synuclein positive pathology associated with PD i.e. *C9orf72* expansions do not cause PD but can cause parkinsonism by affecting the same neuronal population.

2.5 C9orf72 expansions and multiple sclerosis

We (**paper 1**) also noticed an increased incidence of demyelinating disease in *C9orf72*-ALS probands and their families (Cooper-Knock *et al.* 2012). To further investigate this, in **paper 4**, a cohort of multiple sclerosis (MS) cases was screened but no *C9orf72* expansions were identified. However, also in **paper 4**, we described a small number of prospectively identified cases with MS who subsequently developed ALS; amongst these individuals there was a significantly higher than expected number of *C9orf72* expansions (Ismail *et al.* 2013). We suggest that, rather than *C9orf72* expansions causing MS, perhaps MS increases the penetrance of the *C9orf72* expansion. This was supported by the fact that *C9orf72*-related ALS was more rapidly progressive in the patients with a previous history of MS.

2.6 Other phenotypes associated with C9orf72-disease

C9orf72 has also been screened for in Huntington disease (HD)-like syndromes, where repeat expansions were found in 7 cases, at a frequency of 1.7% (Beck *et al.* 2013). Subsequently, in a large cohort of 514 HD phenocopies, expansions were identified in 2% of cases thereby establishing *C9orf72* repeat expansions as the most commonly identified genetic cause of a HD mimic syndrome (Hensman Moss *et al.* 2014). Rare cases of corticobasal syndrome have been reported as having a *C9orf72* expansion (Lesage *et al.* 2013, Lindquist *et al.* 2013) as has a case of progressive supranuclear palsy (Lesage *et al.* 2013). In a cohort of 209

spinocerebellar ataxia cases, an expansion was found a single case, whose father also carried the repeat and had ALS (Fogel *et al.* 2012). In addition, two siblings have been reported with the C9orf72 expansion, one of whom developed ALS but the other developed multiple system atrophy (Goldman *et al.* 2014). Finally, a *C9orf72* expansion was also found in a Finnish patient with a dysplastic gangliocytoma; this individual also exhibited characteristic *C9orf72*-neuropathology (Ferrari *et al.* 2014).

Since *C9orf72* repeat expansions are associated with dementia, cases with AD were screened to determine whether the GGGGCC-repeat also contributed to AD. Some studies have reported *C9orf72* expansions at frequencies of <1% but notably this includes cases with pathologically confirmed AD (Cacace *et al.* 2013, Harms *et al.* 2013, Kohli *et al.* 2013). However, other reports have failed to identify any expansions of >30 repeats in AD cases (Majounie *et al.* 2012, Sha *et al.* 2012, Ticozzi *et al.* 2013). Cohorts of schizophrenia patients have also been screened for the expansion (Huey *et al.* 2013, Fahey *et al.* 2014); however, only one report has identified any expansions and only in <1% of schizophrenic patients (Galimberti *et al.* 2014). It remains to be established if C9orf72 expansions are truly causative of many of these phenotypes or whether the small numbers of cases identified represent chance associations.

List of Papers:

Paper 1:

Cooper-Knock J+, Hewitt C+, Highley JR, Brockington A, Milano A, Man S, Martindale J, Hartley J , Walsh T, Gelsthorpe C, Baxter L, Forster G, Fox M, Mok K, McDermott CJ, Traynor B, Kirby J, Hardy J, Wharton SB, Ince PG, Shaw PJ. Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. Brain. 2012;135(Pt 3):751-64.

- Joint first author

- Extensive role in collecion of samples, performing genetic screening and data analysis

- Primary role in clinical characterisation

- Drafted sections of the manuscriptrelating to clinical characterisation

Paper 2:

Majounie E, Renton AE, Mok K, Nicalou N, Waite A, Rollinson S, Chiò A, Restagno G, Simon-Sanchez J, van Swieten J, Abramzon Y, Johnson JO, Sendtner M, Pamphlett R,. Orrell RW, Mead S, Houlden H, Rohrer JD, Morrison K, Talbot K, Ansorge O, The Chromosome 9-ALS/FTD Consortium (including **Cooper-Knock J**), The ITALSGEN Consortium, Englund E, Borghero G, McCluskey L, Trojanowski JQ, van Deerlin VM, Schellenberg GD, Nalls GA, Drory V, Brice A, Drepper C , Williams N, Kirby J, Shaw P, Hardy J, Singleton A, Tienari PJ, Heutink P, Morris H, Pickering-Brown A, Traynor BJ Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. **Lancet Neurol.** 2012 Apr;11(4):323-30.

- Responsible for collecting samples and clinical data from Sheffield cohort

Paper 3:

Cooper-Knock, J., Frolov A., HighleyJ.R., Charlesworth G., Kirby J., Milano A.*, et al.* (2013). "*C9ORF72* expansions, parkinsonism, and Parkinson disease: a clinicopathologic study." Neurology **81**(9): 808-811.

- Joint first author

- Concieved study, collected samples and conducted data analysis

- Drafted all sections of manuscript

Paper 4:

Ismail, A., **Cooper-Knock J.**, Highley J.R., Milano A., Kirby J., Goodall E., et al. (2013). "Concurrence of multiple sclerosis and amyotrophic lateral sclerosis in patients with hexanucleotide repeat expansions of *C9ORF72*." J Neurol Neurosurg Psychiatry 84(1): 79-87.

- Joint first author

- Collected samples for genetic screening and conducted data analysis

- Performed transcriptome work and validation work in CSF

- Drafted sections relating to genetic screening, microarray analysis and CSF work

doi:10.1093/brain/awr365 Brain 2012: 135; 751–764 | 751

BRAIN
A JOURNAL OF NEUROLOGY

Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72

Johnathan Cooper-Knock,[1,*] Christopher Hewitt,[1,*] J. Robin Highley,[1,*] Alice Brockington,[1] Antonio Milano,[2] Somai Man,[2] Joanne Martindale,[2] Judith Hartley,[1] Theresa Walsh,[1] Catherine Gelsthorpe,[1] Lynne Baxter,[1] Gillian Forster,[1] Melanie Fox,[1] Joanna Bury,[1] Kin Mok,[3] Christopher J. McDermott,[1] Bryan J. Traynor,[4,5] Janine Kirby,[1] Stephen B. Wharton,[1] Paul G. Ince,[1] John Hardy[3] and Pamela J. Shaw[1]

1 Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, 385A Glossop Road, Sheffield S10 2HQ, UK
2 Sheffield Diagnostic Genetic Service, Sheffield Children's NHS Foundation Trust, Western Bank, Sheffield S10 2TH, UK
3 Department of Molecular Neuroscience and Reta Lila Weston Laboratories, Institute of Neurology, University College London, Queen Square House, London WC1N IPJ, UK
4 Neuromuscular Diseases Research Unit, Laboratory of Neurogenetics, National Institute on Ageing, National Institutes of Health, Bethesda, MD 20892, USA
5 Department of Neurology, Brain Sciences Institute, Johns Hopkins University, Baltimore, MD 21287, USA

*These authors contributed equally to this work.

Correspondence to: Professor P. J. Shaw,
Sheffield Institute for Translational Neuroscience (SITraN),
University of Sheffield, 385A Glossop Road,
Sheffield S10 2HQ,
UK
E-mail: pamela.shaw@sheffield.ac.uk

Intronic expansion of the GGGGCC hexanucleotide repeat within the *C9ORF72* gene causes frontotemporal dementia and amyotrophic lateral sclerosis/motor neuron disease in both familial and sporadic cases. Initial reports indicate that this variant within the frontotemporal dementia/amyotrophic lateral sclerosis spectrum is associated with transactive response DNA binding protein (TDP-43) proteinopathy. The amyotrophic lateral sclerosis/motor neuron disease phenotype is not yet well characterized. We report the clinical and pathological phenotypes associated with pathogenic *C9ORF72* mutations in a cohort of 563 cases from Northern England, including 63 with a family history of amyotrophic lateral sclerosis. One hundred and fifty-eight cases from the cohort (21 familial, 137 sporadic) were post-mortem brain and spinal cord donors. We screened DNA for the *C9ORF72* mutation, reviewed clinical case histories and undertook pathological evaluation of brain and spinal cord. Control DNA samples ($n = 361$) from the same population were also screened. The *C9ORF72* intronic expansion was present in 62 cases [11% of the cohort; 27/63 (43%) familial, 35/500 (7%) cases with sporadic amyotrophic lateral sclerosis/motor neuron disease]. Disease duration was significantly shorter in cases with *C9ORF72*-related amyotrophic lateral sclerosis (30.5 months) compared with non-*C9ORF72* amyotrophic lateral sclerosis/motor neuron disease (36.3 months, $P < 0.05$). *C9ORF72* cases included both limb and bulbar onset disease and all cases showed combined upper and lower motor neuron degeneration (amyotrophic lateral sclerosis). Thus, clinically, *C9ORF72* cases show the features of a relatively rapidly progressive, but otherwise typical, variant of amyotrophic lateral sclerosis associated with both familial and sporadic presentations. Dementia was present in the patient or a close family member in 22/62 cases with *C9ORF72* mutation (35%) based on diagnoses established from retrospective clinical

Received November 9, 2011. Revised December 20, 2011. Accepted December 21, 2011

case note review that may underestimate significant cognitive changes in late disease. All the *C9ORF72* mutation cases showed classical amyotrophic lateral sclerosis pathology with TDP-43 inclusions in spinal motor neurons. Neuronal cytoplasmic inclusions and glial inclusions positive for p62 immunostaining in non-motor regions were strongly over-represented in the *C9ORF72* cases. Extra-motor pathology in the frontal cortex ($P < 0.0005$) and the hippocampal CA4 subfield neurons ($P < 0.0005$) discriminated *C9ORF72* cases strongly from the rest of the cohort. Inclusions in CA4 neurons were not present in non-*C9ORF72* cases, indicating that this pathology predicts mutation status.

Keywords: amyotrophic lateral sclerosis; C9ORF72; dementia; neurodegeneration

Abbreviations: ALS = amyotrophic lateral sclerosis; FTD = frontotemporal dementia; FTLD = frontotemporal lobar degeneration; MND = motor neuron disease

Introduction

Amyotrophic lateral sclerosis (ALS) is a rapidly progressive neurodegenerative disorder affecting the motor neurons in the cerebral cortex, brainstem and spinal cord. Progressive destruction of motor neurons leads to a clinical syndrome of muscle weakness, wasting and paralysis resulting in death typically within 2–3 years. Only one drug, riluzole, extends survival and its effects are of modest impact. The understanding of disease pathogenesis is gradually increasing, particularly in relation to genetically determined subtypes of ALS (Ferraiuolo *et al*., 2011), but there is poor understanding of the basic mechanisms of motor neuron injury in sporadic ALS.

ALS affects around six people per 100 000 in the UK with 5–10% of cases having familial disease, usually with autosomal dominant inheritance. Onset is usually in the 6th or 7th decade, although familial cases frequently have a younger age of onset. Mutations in a growing number of genes including superoxide dismutase 1 (*SOD1*), TAR DNA binding protein (*TARDBP*), fused in sarcoma (*FUS*), valosin-containing protein (*VCP*), factor-induced gene 4 (*FIG4*), angiogenin (*ANG*), ubiquilin 2 (*UBQLN2*) and optineurin (*OPTN*) have been shown to be causative in ~30% of adult-onset familial ALS and in a smaller proportion of sporadic ALS cases (Ticozzi *et al*., 2011). In addition, genome-wide association studies have identified variants in several other genes, including ataxin 2 (*ATXN2*) and *UNC13A*, that are associated with increased risk of developing sporadic ALS (Lambrechts *et al*., 2003; van Es *et al*., 2009; Elden *et al*., 2010). Understanding how variations in these genes cause motor neuron degeneration is key to improving our understanding of disease pathophysiology and to the development of more powerful neuroprotective therapies.

In addition to the genes described earlier, genetic linkage to several other genomic regions, including 9p21 (ALS-FTD), 18q21 (ALS3) and 20p13 (ALS7), has been demonstrated in familial ALS cases (Ticozzi *et al*., 2011). The chromosome 9p21 locus has been intensively investigated in recent years by researchers interested both in ALS and the related condition frontotemporal dementia (FTD). This locus was first described in families with a high proportion of concurrent ALS and FTD (Hosler *et al*., 2000) and the locus was more recently refined in ALS cases to a 20 single nucleotide polymorphism common haplotype spanning a 140-kb segment (Mok *et al*., 2011). Linkage to 9p21 has generated particular interest because it provides strong evidence for the association between ALS and frontotemporal lobar degeneration (FTLD), which share common neuropathological features and have significant clinical overlap (Fecto and Siddique, 2011). In addition, the 9p21 locus was estimated to account for nearly half of familial ALS cases and one-fifth of sporadic ALS cases in the Finnish population (Laaksovirta *et al*., 2010). However, direct sequencing of the coding regions of the three genes at this locus (*MOBKL2B*, *C9ORF72* and *IFNK*) did not reveal any pathogenic variants.

Two groups have now separately identified the gene associated with the 9p21 linkage as a GGGGCC hexanucleotide repeat expansion in intron 1 of the gene *C9ORF72* (NM018325.2) (DeJesus-Hernandez *et al*., 2011; Renton *et al*., 2011). Following demonstration of segregation with ALS in two families, Renton *et al*. (2011) subsequently found the expansion in 46.4% of familial and 21% of sporadic ALS cases in the Finnish cohort and 38.1% of familial ALS cases from USA, Italy and Germany. DeJesus-Hernandez *et al*. (2011) similarly demonstrated segregation of the *C9ORF72* expansion in a large autosomal dominant ALS/FTD kindred (VSM-20). Subsequent screening of a larger cohort of FTD and ALS cases identified the expansion in 11.7% of familial and 3% of sporadic FTD cases, and 23.5% of familial and 4.1% of sporadic ALS cases. The length of the expansion varied between 6.5–12 kb (700–1600 repeats), while the majority of controls contained two repeats.

Neuropathology in both FTLD and ALS cases with *C9ORF72* expansions showed TDP-43-positive neuronal and glial inclusions and a higher proportion of nuclear RNA foci in frontal cortex and spinal cord neurons. No unique clinical phenotype was associated with this subtype of ALS or FTD.

We sought to confirm the relevance of this important finding in a large cohort of patients with ALS/motor neuron disease (MND) from Northern England, and to investigate clinical and pathological differences between cases with and without the repeat expansion. These patients represent a well-characterized cohort of 563 ALS cases, including 63 familial ALS index cases, with serial clinical assessment performed every 2–3 months throughout the disease course. Post-mortem pathological evaluation was available in 28% of these cases.

Materials and methods

Cases with amyotrophic lateral sclerosis and controls

DNA was extracted from 42 familial ALS index cases and 363 patients with sporadic ALS from the Sheffield MND Blood DNA Biobank. Additional DNA samples were isolated from 21 familial ALS index cases and 137 sporadic ALS cases in the Sheffield Brain Tissue Bank. In total, 563 ALS cases were screened: 63 familial ALS index cases and 500 sporadic ALS cases. Familial ALS cases were defined as individuals with one or more first or second degree relatives with a confirmed diagnosis of ALS. All 563 cases were reviewed by a senior consultant neurologist (C.J.M. or P.J.S.) and diagnosed with definite or probable ALS, as defined by the El Escorial criteria (Brooks *et al*., 2000). A full family history was taken from each patient. Patients with known mutations in *SOD1* ($n = 14$), *TARDBP* ($n = 5$), *FUS* ($n = 4$), *ANG* ($n = 1$), *OPTN* ($n = 1$), charged multi-vesicular protein 2B (*CHMP2B*) ($n = 4$) and vesicle associated membrane protein 2B (*VAPB*) ($n = 1$) were included in the *C9ORF72* screening. DNA was extracted from blood using the Nucleon™ Blood and Cell Culture Genomic Extraction kit (Tepnel) according to the manufacturer's protocol, while DNA was extracted from fresh frozen cerebellar samples using the Soft Tissue DNA Extraction Kit (Tepnel). Control DNA ($n = 361$) was extracted from blood donated by partners or unrelated carers of patients with ALS. All samples were from UK Caucasians. The South Sheffield Research Ethics Committee approved the study, and informed consent was obtained for all samples.

Determination of the clinical phenotype of patients with ALS with the C9ORF72 hexanucleotide expansion

Clinical notes of patients found to carry the *C9ORF72* expansion were reviewed in a systematic fashion to identify details of the disease phenotype including gender, age of onset, disease duration, disease variant, details of family history and the presence of any cognitive impairment. It should be noted that patients in this cohort only underwent formal cognitive evaluation when a clinical problem was identified and did not undergo routine serial neuropsychological evaluation during the disease course. Therefore, the recorded incidence of frontotemporal dysfunction in this cohort is likely to be an underestimate, as prospective neuropsychological studies show a higher prevalence of frontal lobe dysfunction in patients with ALS (Phukan *et al*., 2011).

Screening for the C9ORF72 hexanucleotide repeat sequence by repeat primed polymerase chain reaction

Genomic DNA (100 ng) was amplified using the primers and method described by Renton *et al*. (2011) with a minor adjustment to the primer ratio: Forward:Reverse:Anchor = 8:1:8. Detailed methodology is provided in the Supplementary Material. Fragments were analysed on an ABI3730 capillary analyser (Applied Biosystems, Life Technologies Corporation) using a 60-s injection time. Fragment data were analysed using Peak Scanner Software (Applied Biosystems, Life Technologies Corporation).

Neuropathological evaluation

The brain and spinal cord tissues were donated to the Sheffield Brain Tissue Bank for research, with the consent of the next of kin. The donation procedure and use of the tissue in this project were undertaken with research ethical committee approval. Tissue was available for detailed pathological evaluation from 19 of the 22 brain tissue bank cases with the hexanucleotide expansion of *C9ORF72*. These cases were compared with up to 96 (see below and Table 3) ALS cases, which were negative for the *C9ORF72* expansion and three neurologically normal controls. For some cases, one cerebral hemisphere, half the midbrain and brainstem, a portion of the cerebellum and segments of the spinal cord at various levels were rapidly frozen in liquid nitrogen at autopsy and stored at -80°C. The remainder of the CNS was formalin-fixed. For the other cases, only a portion of the cerebellum was frozen, and the whole brain and spinal cord were formalin fixed. Selected blocks (including lumbar, thoracic and cervical spinal cord, medulla, midbrain, hippocampus, and frontal, temporal and motor neocortex) were processed to paraffin.

In addition to routine neurohistology with tinctorial preparations, immunohistochemistry was performed for p62/sequestosome 1, TAR DNA-binding protein 43 (TDP-43), FUS, OPTN, CD68 and C9ORF72 (Supplementary Table 1) where paraffin tissue was available. The latter antibody is commercially available (Santa Cruz Biotechnology Inc.) and was selected on the basis that it labelled a protein of the molecular weight of C9ORF72 on western blotting (data not shown). Immunohistochemistry was performed on all available cases with the hexanucleotide expansion as well as three cases with sporadic ALS without the expansion and three neurologically healthy controls.

To characterize the distribution of pathology in cases with ALS in the Sheffield Brain Tissue Bank, the extent of p62-positive pathology was assessed. In all regions assessed, a single 6-μm section was examined and the number of neuronal cytoplasmic inclusions in the region of interest assessed semi-quantitatively as 0–4 (low), 5–9 (intermediate) or 10 or more (high). The regions of interest assessed were the anterior horn of the spinal cord, at mid-cervical and lumbar levels; the hypoglossal, dorsal vagal and ambiguus nuclei of the medulla; the dentate granule cell layer and CA4 subregion of the hippocampus and the frontal and motor cortices. In the cortical regions, quantification of neuronal cytoplasmic inclusions was carried out in 10 fields ($\times 25$ objective).

In addition, the number of neuronal cytoplasmic inclusions in the CA4 subregion of the hippocampus was assessed using the same semi-quantitative scheme on sections that had been immunostained for TDP-43 and OPTN.

Statistical methods

Differences between phenotypic characteristics of groups were determined by an independent-samples *t*-test. Differences between gender ratios and site of onset and differences in the ubiquitylated neuronal cytoplasmic inclusion load were calculated using a chi-squared (χ^2) test. A significance level of $P < 0.05$ was used for all tests.

Results

Genetic screening

ALS-associated pathological expansions in *C9ORF72* have been defined as >30 repeats, whereas the majority of controls have

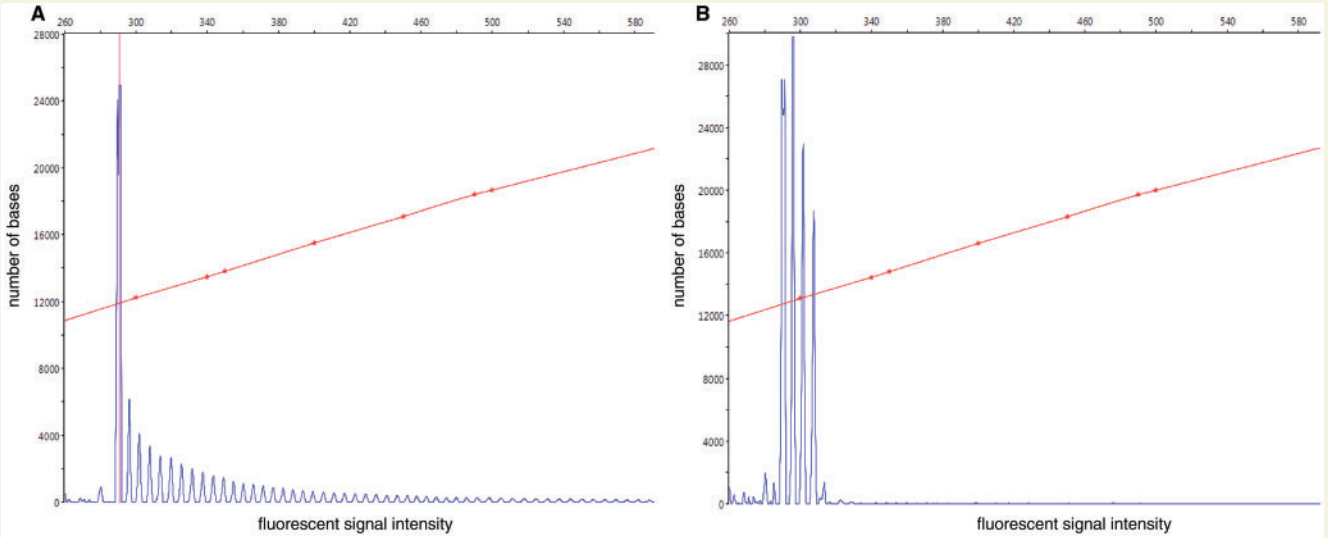


Figure 1 Graphical representation of fragments generated in repeat primed polymerase chain reaction displayed using Peak Scanner Software. Individual peaks represent fragments of increasing length at 6 bp intervals, corresponding to single GGGGCC repeats; 280 bp is minimum polymerase chain reaction product length. (**A**) Graph from a patient with ALS with a pathological expansion in *C9ORF72* showing exponentially tailing series of peaks consistent with a repeat length >30. (**B**) Graph from a patient with ALS without pathological expansion showing repeat length of approximately four repeats.

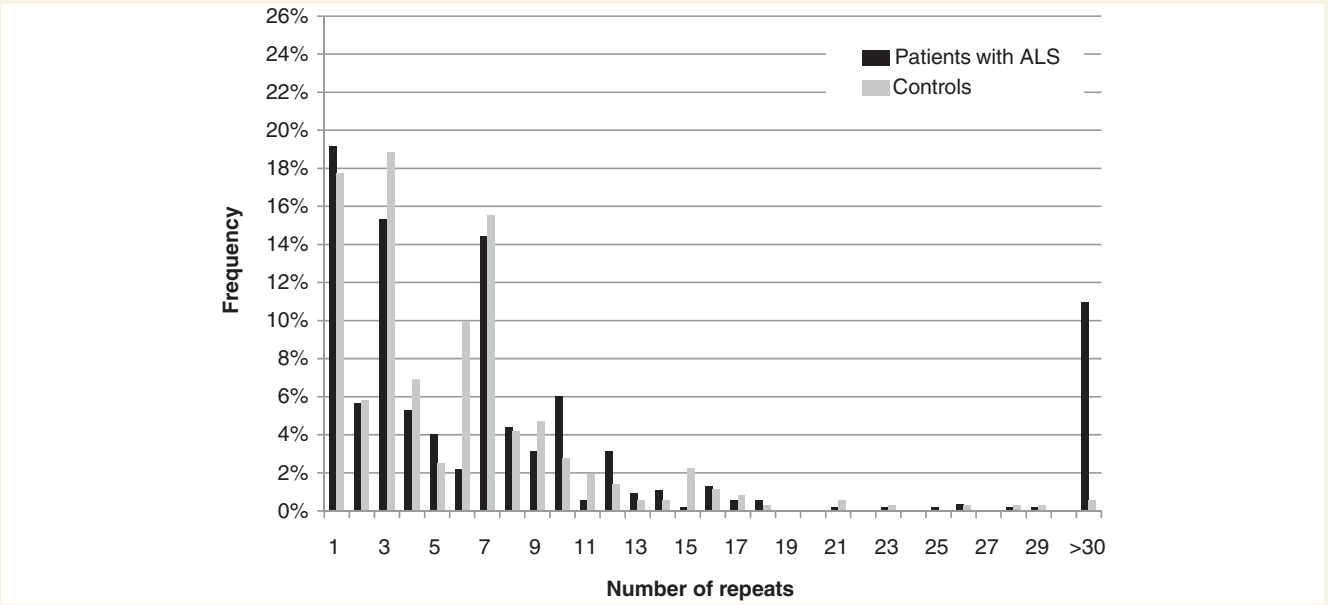


Figure 2 Graph showing the lengths of the GGGGCC hexanucleotide repeats in ALS cases ($n = 563$) and controls ($n = 361$); 62/563 (11%) of cases with ALS and 2/361 (0.6%) of control cases had repeat lengths of >30.

⩽3 repeats (DeJesus-Hernandez *et al.*, 2011; Renton *et al.*, 2011). In our cohort of 563 ALS cases, 62 (11%) ALS cases (27 familial ALS, 35 sporadic ALS) were found to have expansions >30 repeats (Fig. 1). The median number of repeats in ALS cases without pathological expansions was four compared with five in controls (Fig. 2). The length of a pathologically expanded repeat (>30 repeats) cannot be accurately quantified by repeat primed polymerase chain reaction. This technique is only sufficient to segregate individuals with large expansions >30 from those without. An expanded repeat >30 was detected in two control cases (2/361 = 0.6%), a 76-year-old male and a 46-year-old female, neither of whom had any relevant past medical or family history.

Pathological expansions accounted for 27/63 (43%) of our familial ALS index cases. Analysis of family trees of all index cases revealed 13 families with a clear autosomal dominant pattern of inheritance and we were able to demonstrate segregation

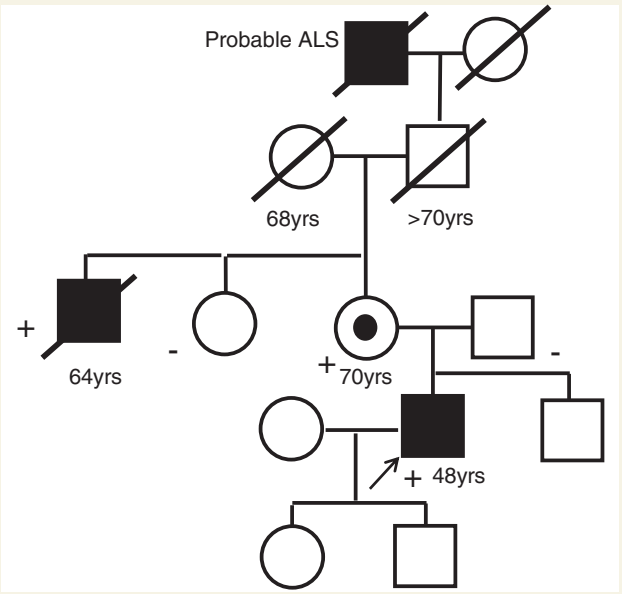


Figure 3 Family tree from a familial ALS case with the hexanucelotide repeat sequence in *C9ORF72*. The expansion segregated with disease in two cases and was present in an obligate carrier. + = carrier of expansion, − = confirmed absence of the expansion. Current age or age of death is shown where information was available.

of the expanded allele in one family with evidence of incomplete penetrance (Fig. 3). In familial ALS cases with previously identified mutations in *SOD1*, *TARDBP*, *FUS*, *ANG*, *CHMP2B* and *VAPB*, an expanded *C9ORF72* repeat >30 was identified in a case with an p.Ala321Val mutation in *TARDBP* (Kirby *et al*., 2010) and also in a case with a benign p.Gly174del alteration in *FUS* reported previously (Hewitt *et al*., 2010). An unreported case with a p.Glu322Lys *OPTN* substitution also showed an expansion.

Pathological expansions were also present in 35/500 (7%) of our cases with sporadic ALS. However, for 16/35 cases there was either a family history of dementia or other neurological or neuromuscular disease; or at least one parent died before the age of 70 years. It is therefore likely that a proportion of these cases represent unrecognized familial disease. For the remaining 19/35 sporadic ALS cases with the expansion, a complete family history was available and in these cases there was no family history of neurological disease and both parents lived into late old age. Therefore, it is likely that these cases represent true sporadic disease, although incomplete penetrance cannot be excluded. DNA from parents of these apparently sporadic cases was not available for analysis to confirm the occurrence of *de novo* mutations.

Clinical features of ALS cases with pathological expansions in C9ORF72

The clinical features of the cases with pathological expansions in *C9ORF72* are summarized in Table 1. The mean age of onset was 57.3 years (range 27–74 years). Of these, 60% of cases had limb onset disease, 31% had bulbar onset disease; 6% had multi-focal disease onset, one patient presented with dementia and for one patient the time of onset was not known. The mean disease duration from symptom onset in these patients was 30.5 months (range 7–60 months); three patients are alive at the time of writing and for one patient survival information was not available.

Clinical characteristics of patients with the *C9ORF72* expansion were compared with the remainder of the screened cohort (Table 2). The *C9ORF72* patients had a significantly lower age at onset than non-*C9ORF72* cases (mean age of onset 57.3 years, SD 8.9 years compared with mean 60.1 years, SD 12.3 years; $P = 0.03$, d.f. = 494, $t = 2.20$), but not significantly different to the overall familial ALS cohort ($P = 0.43$, d.f. = 107, $t = 0.78$). Likewise, patients with the *C9ORF72* expansion had a significantly shorter duration of disease than non-*C9ORF72* cases [mean (SD) duration of disease 30.5 (13.3 months) compared with mean (SD) 36.3 (28.5 months); $P = 0.01$, d.f. = 397, $t = 2.44$] but not significantly different to the overall familial ALS cohort ($P = 0.32$, d.f. = 100, $t = 1.00$). The subgroup of patients with sporadic ALS with the *C9ORF72* expansion differed from patients with familial ALS with the expansion, in that they had a significantly higher age of onset [mean 59.5 (SD 7.0 years) in sporadic ALS, 54.3 (SD 10.4 years) in familial ALS, $P = 0.03$, d.f. = 59, $t = 2.21$], but no difference in age of onset to the screened cohort overall. There was no difference between patients with sporadic ALS and patients with familial ALS with the expansion, with respect to duration of disease [mean (SD) 28.6 (12.0 months) in sporadic ALS, and 33.3 (14.8 months) in familial ALS, $P = 0.21$, d.f. = 56, $t = 1.26$].

In our cohort of patients with the *C9ORF72* expansion, 5/27 (19%) familial ALS cases and 5/35 (14%) sporadic ALS cases had evidence on clinical and neuropsychological testing of FTD. An additional 12 cases had a family history of dementia in first- or second-degree relatives, six of whom had familial ALS and six of whom had sporadic ALS. Overall 22/62 (35%) of patients with the expansion had either a personal diagnosis of dementia or a family history of dementia in first- or second-degree relatives. In this cohort, routine neuropsychological assessment was not performed in the absence of a clinically apparent cognitive problem, so subclinical cognitive dysfunction was not evaluated.

Several of the patients carrying the hexanucleotide expansion in *C9ORF72* were either diagnosed with or had a family history of other non-dementia neurological or neuromuscular disease, particularly neurodegenerative disease: four patients had a family history of Parkinson's disease and one patient had a comorbid diagnosis of Parkinson's disease that was confirmed at post-mortem; thus 4/62 (6.5%) patients had either a diagnosis of Parkinson's disease or a family history of Parkinson's disease. Two patients had evidence of demyelinating disease and one further patient had a family history of multiple sclerosis; thus 3/62 (5%) patients had either a diagnosis or a family history of demyelinating disease. Other noteworthy findings included a patient with a family history of ALS, Charcot–Marie–Tooth disease and dementia; one patient with early onset cataracts which also occurred in his mother; one patient with a family history of Huntington's disease and one patient with a brother who died with a diagnosis of muscular dystrophy.

Table 1 Summary of phenotypic information from patients with the hexanucleotide repeat expansion of *C9ORF72*

Patient	Sex	Age at onset (years)	Duration of disease (months)	Site of onset	Variant of disease	Family history	Cognitive impairment	Family history and other noteworthy features
Brain Tissue Bank								
1	Male	69	26	Limb	ALS	Familial	None	Autosomal dominant pattern of familial ALS
2	Female	59	40	Limb	ALS	Familial	None	Autosomal dominant pattern of familial ALS
3	Male	42	50	Limb	ALS	Familial	None	Autosomal dominant pattern of familial ALS
4	Male	66	14	Bulbar	ALS	Familial	NA	Autosomal dominant pattern of familial ALS
5	Male	64	31	Limb	ALS	Familial	Undefined dementia	Nephew is Patient 27.
6	Female	62	24	Bulbar	ALS	Familial	None	Autosomal dominant pattern of familial ALS. Patient diagnosed with multiple sclerosis. Benign polymorphism Gly174del in FUS/TLS.
7	Female	56	43	Limb	ALS	Familial	Diagnosed FTD	Maternal family history of ALS, Charcot–Marie–Tooth and early onset undefined dementia.
8	Female	50	28	Bulbar	ALS	Familial	None	Autosomal dominant pattern of familial ALS. Patient carries a p.Glu322Lys substitution in *OPTN*
9	Male	47	19	Multi-focal	ALS	Familial	None	Autosomal dominant pattern of familial ALS. Patient and his father diagnosed with Parkinson's disease.
10	Female	63	43	Cognitive	ALS	Familial	Diagnosed FTD	Mother diagnosed with early onset undefined dementia. Brother diagnosed with ALS.
11	Female	61	58	Limb	ALS	Familial	None	Two sisters diagnosed with ALS
12	Female	65	12	Bulbar	ALS	Sporadic	None	
13	Female	67	26	Limb	NA	Sporadic	None	
14	Male	63	11	Limb	ALS	Sporadic	None	
15	Male	56	13	Bulbar	ALS	Sporadic	None	
16	Female	61	40	Bulbar	ALS	Sporadic	Undefined dementia	
17	Female	58	7	Limb	ALS	Sporadic	None	
18	Female	61	42	Limb	ALS	Sporadic	None	
19	Male	62	20	Bulbar	ALS	Sporadic	None	
20	Male	45	14	Limb	ALS	Sporadic	None	
21	Female	51	38	Multifocal	ALS	Sporadic	None	
Blood DNA Biobank								
22	Female	47	18	Limb	ALS	Familial	Diagnosed FTD	Autosomal dominant pattern of familial ALS
23	Male	61	38	Limb	ALS	Familial	None	Brother and sister diagnosed with ALS. Father diagnosed with Huntington's disease.
24	Female	27	14	Bulbar	ALS	Familial	None	Autosomal dominant pattern of familial ALS. Maternal family history of AD; paternal family history of undefined dementia.
25	Female	50	18	Bulbar	ALS	Familial	Undefined dementia	Father diagnosed with FTD-ALS. Autosomal dominant pattern of familial ALS
26	Female	48	52	Limb	ALS	Familial	None	Brother diagnosed with ALS. Maternal aunt diagnosed with Alzheimer's disease. Patient notably athletic.
27	Male	48	ALIVE	Bulbar	ALS	Familial	None	Maternal uncle is Patient 5. Maternal grandmother diagnosed with early onset dementia.
28	Female	61	22	Limb	ALS	Familial	None	Autosomal dominant pattern of familial ALS
29	Female	44	43	Bulbar	ALS	Familial	None	Autosomal dominant pattern of familial ALS
30	Female	61	43	Bulbar	NA	Familial	None	Paternal uncle diagnosed with ALS. Father diagnosed with Parkinson's disease.
31	Female	64	ALIVE	Limb	ALS	Familial	None	Brother diagnosed with ALS. Paternal grandfather diagnosed with undefined dementia.
32	Female	45	34	Bulbar	ALS	Familial	NA	Autosomal dominant pattern of familial ALS. Mother diagnosed with FTD-ALS. Maternal grandmother diagnosed with multiple sclerosis.
33	Male	51	17	Limb	ALS	Familial	None	Sister diagnosed with FTD-ALS. Patient and mother diagnosed with early onset cataracts.
34	Male	65	52	Limb	ALS	Familial	None	Nephew diagnosed with ALS. Patient notably athletic
35	Male	63	13	Multifocal	ALS	Familial	None	Paternal aunt diagnosed with ALS
36	Female	NA	NA	NA	NA	Familial	NA	
37	Female	37	58	Limb	ALS	Familial	None	Mother probable ALS. Patient has pAla321Val mutation in *TARDBP*
38	Male	56	24	Limb	ALS	Sporadic	None	

(continued)

Table 1 Continued

Patient	Sex	Age at onset (years)	Duration of disease (months)	Site of onset	Variant of disease	Family history	Cognitive impairment	Family history and other noteworthy features
39	Male	56	41	Limb	ALS	Sporadic	None	
40	Female	66	32	Bulbar	ALS	Sporadic	None	Father died of undefined dementia
41	Male	60	35	Bulbar	ALS	Sporadic	Undefined dementia	Mother diagnosed with Alzheimer's disease
42	Female	50	27	Limb	ALS	Sporadic	None	Father diagnosed with early onset dementia.
43	Male	60	32	Limb	ALS	Sporadic	None	Mother diagnosed with early onset dementia. Patient notably athletic.
44	Male	43	40	Limb	ALS	Sporadic	None	Patient notably athletic.
45	Male	64	28	Limb	ALS	Sporadic	None	
46	Male	74	36	Limb	ALS	Sporadic	None	Father and sister diagnosed with undefined dementia.
47	Male	58	60	Limb	ALS	Sporadic	None	Onset coincided with carpal tunnel syndrome therefore difficult to determine exact date.
48	Female	57	20	Limb	ALS	Sporadic	None	Demyelination noted on MRI of CNS.
49	Female	62	24	Limb	ALS	Sporadic	None	Maternal aunt diagnosed with Alzheimer's disease. Patient notably athletic
50	Female	57	21	Multifocal	ALS	Sporadic	None	
51	Male	60	27	Bulbar	ALS	Sporadic	None	
52	Male	63	ALIVE	Limb	ALS	Sporadic	Undefined dementia	Brother and father diagnosed with schizophrenia
53	Male	61	24	Limb	ALS	Sporadic	None	Mother diagnosed early onset Alzheimer's disease and Parkinson's disease. Patient notably athletic.
54	Male	51	21	Bulbar	ALS	Sporadic	None	Previous poliomyelitis which left him with wasted right leg
55	Male	50	27	Bulbar	ALS	Sporadic	None	
56	Female	59	22	Limb	ALS	Sporadic	Undefined dementia	
57	Female	71	24	Limb	ALS	Sporadic	Diagnosed FTD	
58	Female	65	40	Limb	ALS	Sporadic	None	Patient notably athletic.
59	Female	52	24	Limb	ALS	Sporadic	NA	
60	Female	63	28	Bulbar	ALS	Sporadic	NA	
61	Female	71	57	Limb	ALS	Sporadic	None	Cousin diagnosed with Parkinson's disease
62	Male	65	36	Limb	ALS	Sporadic	None	Brother suffered muscular dystrophy

NA = data not available.

Table 2 Comparison of phentotypic information from patients the hexanucleotide repeat expansion of *C9ORF72* and the overall screened cohort

Phenotype	*C9ORF72* cases ($n = 62$)	*C9ORF72* sporadic ALS cases ($n = 35$)	*C9ORF72* familial ALS cases ($n = 27$)	Non-*C9ORF72* ALS cases in the screened cohort ($n = 501$)
Mean (SD) age of onset (years)	57.3 (8.9)	59.5 (7.0)	54.3 (10.4)	60.1 (12.3)
Mean (SD) disease duration (months)	30.5 (13.3)	28.6 (12.0)	33.3 (14.8)	36.3 (28.5)
Limb onset (%)	60	66	51	56
Bulbar onset (%)	31	29	34	25
Gender ratio (males:females)	1:1.2	1.1:1	1:1.7	1.3:1

The patients with *C9ORF72* expansion had limb onset disease in 37/62 (60%) cases and bulbar onset disease in 19/62 (31% cases) (Table 2). These proportions are similar to those for the cohort without the *C9ORF72* expansion (56 and 25%, respectively). There was no significant difference in gender ratio ($P = 0.09$, d.f. = 1, $\chi^2 = 2.85$) or site of onset ($P = 0.74$, d.f. = 1, $\chi^2 = 0.10$) between the two groups although the patients with familial ALS with the expansion had a female preponderance (ratio males:females = 1:1.7), which was not present in the patients with sporadic ALS with the expansion (males: females = 1.1:1) or the screened cohort overall (males:females = 1.3:1).

Neuropathological features

All cases with the hexanucleotide repeat expansion of *C9ORF72* showed the classical molecular pathology of ALS (Figs 4 and 5). There was a marked loss of lower motor neurons in the anterior horns of the spinal cord and cranial nerve motor nuclei in the medulla. Bunina bodies were present in some residual motor neurons. All cases showed a moderate to marked microglial reaction on CD68 immunohistochemistry in the pyramidal tract at all levels (white matter underlying motor cortex, mid-crus cerebri,

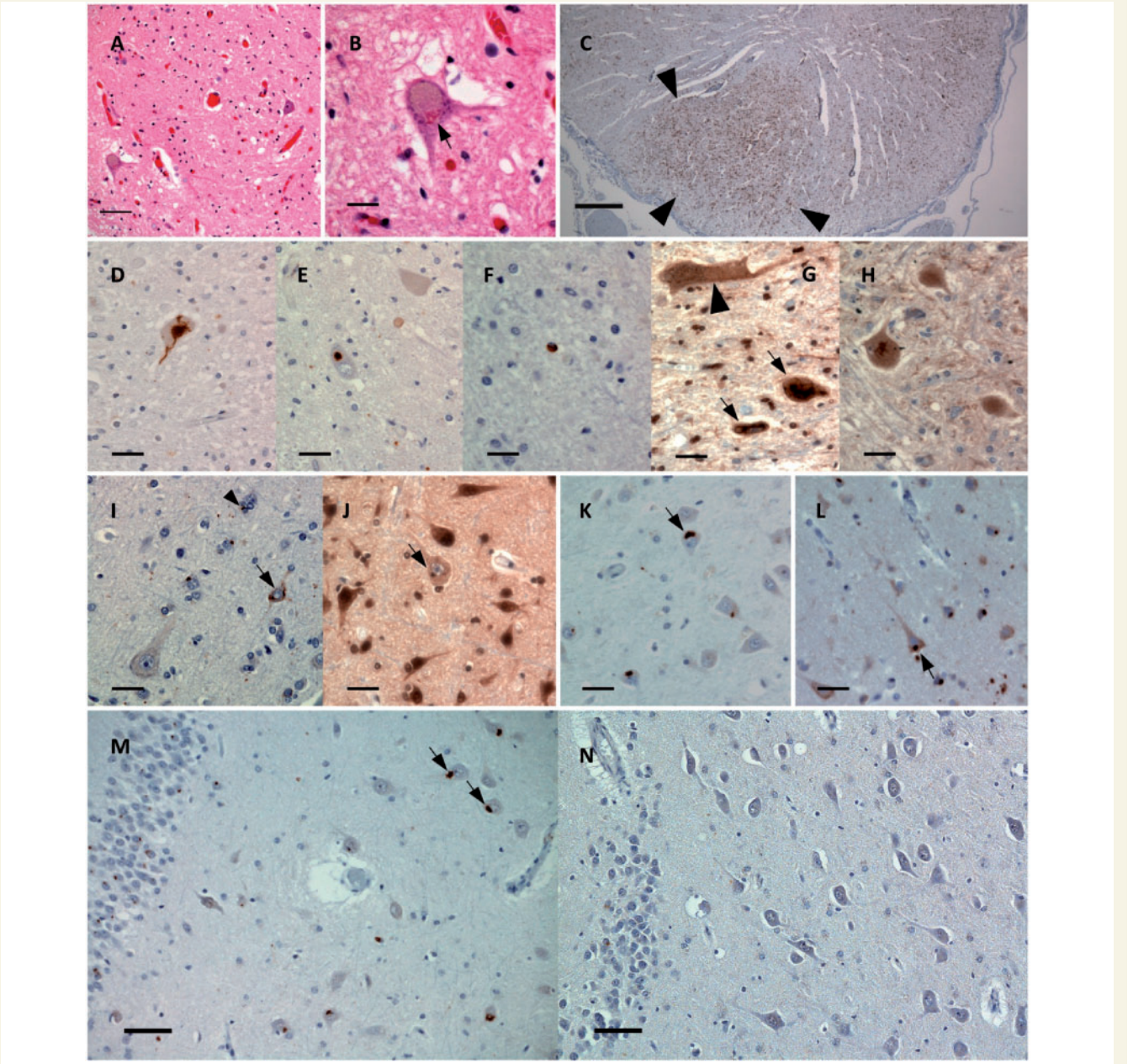


Figure 4 Histological examination of cases with the *C9ORF72* hexanucleotide repeat expansion reveals: depletion of motor neurons from the anterior horns of the spinal cord (**A**) with Bunina bodies (arrow) in residual neurons (**B**); Microglial activation in lateral corticospinal tracts (arrowheads, **C**); ubiquitylated neuronal (skein-like in **D** and compact in **E**) and glial (**F**) cytoplasmic inclusions in the anterior horns of the spinal cord; (**G**) TDP-43 positive skein-like neuronal cytoplasmic inclusions (arrows) and pre-inclusions (arrowhead) in the anterior horns of the spinal cord; (**H**) OPTN positive neuronal cytoplasmic inclusion in the anterior horns of the spinal cord; (**I**) ubiquitylated neuronal (arrow) and glial (arrowhead) cytoplasmic inclusions in the motor cortex; TDP-43 positive neuronal cytoplasmic inclusion in the motor cortex (arrow, **J**); ubiquitylated neuronal cytoplasmic inclusions (arrows) in the CA4 subfield of the hippocampus (**K**) and frontal neocortex (**L**). Lower power view of hippocampal CA4 subfield with adjacent dentate gyrus granule cells (*left*) reveals neuronal cytoplasmic inclusions (arrows) in CA4 of a case of ALS with *C9ORF72* hexanucleotide repeat expansion (**M**) and none in a case without this expansion (**N**). Preparations: haematoxylin and eosin, **A** and **B**; CD68, **C**; p62, **D–F**, **I**, **K–N**; TDP-43, **G** and **J**; optineurin, **H**. Scale bar = 20 μm (**B** and **F**); 30 μm (**D**, **E**, **G** and **H–L**); 60 μm (**A**, **M** and **N**); 500 μm (**C**).

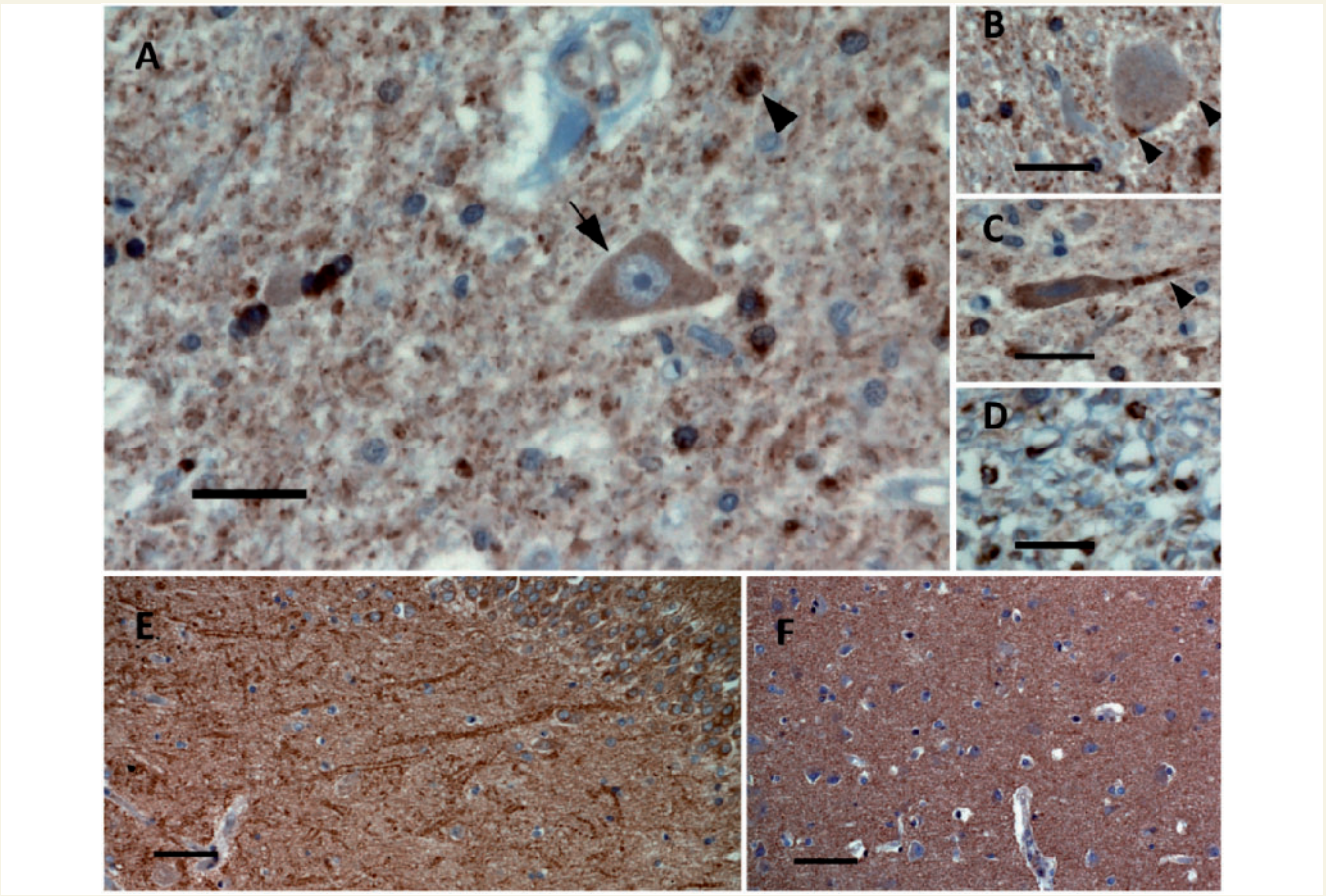


Figure 5 Immunohistochemistry for C9ORF72 shows similar features in neurologically healthy controls and individuals with motor neuron disease with and without the hexanucleotide repeat expansion. The spinal cord stained with C9ORF72 antibody shows: (**A**) Coarse granular staining of the spinal cord anterior horn neuropil with variable cytoplasmic labelling of motor neurons (arrow) and perinuclear labelling of small glial cells (e.g. arrowhead); patches of more intense labelling (arrowheads) on the surfaces of some motor neurons (**B**) that is often more marked in neuronal processes (**C**). In white matter (dorsal column) there is staining around axons corresponding to portions of myelin sheath (**D**). Granular staining of the neuropil with more prominent 'streaks' and only pale neuronal expression is seen in the hippocampus (CA4 subregion with dentate gyrus granule cell layer top right; **E**). Granular neuropil labelling with lesser staining of neurons is also seen in the neocortex (frontal cortex; **F**). Scale bar = 20 μm (**B** and **F**); 30 μm (**A–D**); 60 μm (**E** and **F**).

medullary pyramids, and lateral and anterior corticospinal tracts) indicating corticospinal tract degeneration.

Immunohistochemistry for p62 revealed neuronal and glial cytoplasmic inclusions in the anterior horns of the spinal cord, cranial nerve motor nuclei and motor cortex. These were also detectable using both TDP-43 and OPTN immunohistochemistry. Results of the quantitation of neuronal cytoplasmic inclusions using p62 immunohistochemistry are presented in Table 3. In general terms, there was a significantly higher load of neuronal cytoplasmic inclusions in all regions in individuals with the *C9ORF72* repeat expansion compared with ALS cases without this expansion.

The most striking feature of cases bearing the *C9ORF72* hexanucleotide expansion was the degree of extra-motor p62 positive neuronal cytoplasmic inclusion pathology in contrast to the relative paucity of this finding in cases without the expansion. This feature was most pronounced in the CA4 and CA3 subregions of the hippocampus and, to a lesser extent, the frontal neocortex and dentate gyrus of the hippocampus. To avoid potential bias, hippocampal and frontal neocortex slides prepared for p62 immunohistochemistry from 11 cases with and 11 cases without the expansion were randomly selected and reassessed blind to case identity and *C9ORF72* gene status. For the frontal neocortex, low, intermediate and high levels of neuronal cytoplasmic inclusions were seen in 0, 1 and 10 cases with the expansion and 10, 1 and 0 cases without the expansion, respectively ($\chi^2 = 20$, d.f. = 2, $P < 0.0005$). For the dentate fascia of the hippocampus, low, intermediate and high levels of cytoplasmic inclusions were seen in 3, 1 and 7 cases with the expansion and 9, 0 and 2 cases without the expansion, respectively ($\chi^2 = 6.778$, d.f. = 2, $P = 0.034$). For the CA4 subfield of the hippocampus, high levels of neuronal cytoplasmic inclusions were seen in all 11 cases with the hexanucleotide expansion and low levels in all 11 cases without the expansion ($\chi^2 = 22.000$, d.f. = 1, $P < 0.0005$). The slides were assessed independently by a second neuropathologist who assigned all 22 cases correctly into groups with

Table 3 Numbers of cases with low, intermediate and high levels of neuronal cytoplasmic inclusion pathology in cases with and without the *C9ORF72* hexanucleotide repeat expansion in different regions of the CNS

C9ORF72 repeat expansion	Total number of cases				Per cent			Chi-squared test		
	Low	Intermediate	High	Total	Low	Intermediate	High	χ^2	d.f.	*P*-value
Lumbar cord anterior horns										
Present	3	6	7	16	19	38	44	23.657	2	<0.0005
Absent	32	10	1	43	74	23	2			
Cervical cord anterior horns										
Present	7	5	3	15	47	33	20	6.243	2	0.044
Absent	32	6	2	40	80	15	5			
Medulla motor nuclei										
Present	8	1	6	15	53	7	40	7.800	2	0.02
Absent	12	3	0	15	80	20	0			
Dentate gyrus of hippocampus										
Present	0	0	17	17	0	0	100	36.780	2	<0.0005
Absent	66	8	22	96	69	8	23			
Hippocampus CA4 subfield										
Present	0	0	17	17	0	0	100	67.501	1	<0.0005
Absent	93	0	0	93	100	0	0			
Frontal neocortex										
Present	0	1	18	19	0	5	95	74.495	2	<0.0005
Absent	92	1	1	94	98	1	1			

Table 4 Numbers of cases with *C9ORF72* hexanucleotide repeat showing low, intermediate and high numbers of neuronal cytoplasmic inclusions in the hippocampal dentate gyrus granule cell layer and CA4 subfields using immunohistochemistry for p62, TDP-43 and OPTN

CNS region	p62, *n* (%)	TDP-43, *n* (%)	OPTN, *n* (%)
Dentate gyrus			
Low	0 (0)	10 (58.8)	13 (81.25)
Intermediate	0 (0)	3 (17.6)	2 (12.5)
High	17 (100)	4 (23.5)	1 (6.25)
CA4			
Low	0 (0)	15 (88.2)	4 (25)
Intermediate	0 (0)	1 (5.9)	2 (12.5)
High	17 (100)	1 (5.9)	10 (62.5)

and without the hexanucleotide expansion on the basis of p62-immunoreactive pathology in the CA4 hippocampal subfield. There was thus 100% inter-rater agreement between the assessments of the two observers. Next, the proportion of the neuronal cytoplasmic inclusion pathology seen in the hippocampus that was detected by TDP-43 and OPTN immunohistochemistry was investigated (Table 4). This was markedly less than was seen on p62 immunohistochemistry; in the CA4 subregion, OPTN antibodies detect greater numbers of inclusions than TDP-43 antibodies. The converse is true for the dentate gyrus.

The hippocampal and neocortical p62 labelling took the form of neuronal and glial inclusions in all cortical layers with very few dystrophic neurites, corresponding to Type B according to the harmonized classification system for FTLD-TDP pathology (Mackenzie *et al.*, 2011).

A single case with the hexanucleotide repeat expansion showed combined ALS and multiple sclerosis pathology. This case had previously been demonstrated to also have a benign polymorphism (p.Gly174del) in the gene *FUS* and has been described elsewhere (see Case 4, Hewitt *et al.*, 2010).

Immunohistochemistry for C9ORF72 revealed fine punctate staining throughout the grey matter structures of the CNS. Neuronal cell bodies showed only pale to moderate intensity of staining. There were occasional patches of staining on the cell membrane that was more prominent on axonal hillocks. Neuronal nuclei were negative for C9ORF72. This pattern of staining was reminiscent of that seen in preparations for synaptic markers. There was positive staining of smaller glial cells, mostly having the appearance of oligodendrocytes. In white matter tracts, there was partial circumferential staining around axons, with some

granularity to the staining pattern. In the hippocampal pyramidal cell layer, there was coarse staining of the neuropil that was most intense in the CA4 and CA3 subfields. The intensity of this staining was much less marked in CA2 and CA1. In the hippocampus, the intense C9ORF72 labelling was seen in the same regions that also demonstrated ubiquitylated neuronal cytoplasmic inclusions. However, no difference in the pattern of immunostaining for C9ORF72 was observed between cases with and without the hexanucleotide expansion or in the neurologically healthy control cases.

Discussion

In keeping with the recent reports (DeJesus-Hernandez *et al*., 2011; Renton *et al*., 2011), we find significant variability in the presentation of ALS in patients carrying the *C9ORF72* hexanucleotide expansion so that patients with the expansion manifest the full range of the known ALS phenotype. In addition, we have identified a number of interesting associations between ALS and other neurological diseases in this group of patients. This may suggest that the effect of *C9ORF72* expansions is modified by other genetic or environmental factors to produce variability in the ALS phenotype and even in the area of the nervous system affected. Importantly, we find that with the exception of the granule cells of the dentate fascia, carriers of the *C9ORF72* hexanucleotide repeat expansion account for the extra-motor pathology in our series.

Patterns of inheritance

Hexanucleotide expansions of *C9ORF72* account for a large proportion (43%) of familial ALS cases in our cohort from the North of England and many of these families exhibit clear autosomal dominant inheritance. The high incidence of the *C9ORF72* repeat expansions in apparently sporadic cases is an important observation to explain. It may represent a genuinely inherited pathogenic genotype with variable penetrance or it may represent a high rate of new mutation consistent with the suggestion that the *C9ORF72* gene may be inherently unstable (Renton *et al*., 2011). That incomplete penetrance can occur is demonstrated by a family within our cohort where the expansion is present in the index case and his maternal uncle who are both patients with ALS, and in the mother of the index case who is disease free into old age (Fig. 3). It is noteworthy that for a proportion of the apparently sporadic cases there is a family history of neurological disease; it seems likely that some of these cases represent unrecognized or undisclosed familial disease. In particular, some of the patients with relatives who developed dementia may in fact have familial ALS/FTD. It is interesting that, with regard to age of onset and gender ratio but not disease length, the patients with sporadic ALS with the *C9ORF72* expansion are distinct from the patients with familial ALS with the expansion. The earlier age of onset of familial ALS cases with the mutation may suggest the occurrence of genetic anticipation, although this could not be confirmed in this study as the precise length of the pathological repeat sequences has not yet been determined.

Of note, we have now identified a pathogenic change in 42/63 (67%) of our familial cohort. In two cases, an expansion in *C9ORF72* was found to coexist with another identified putative mutation linked to ALS. That the expansion was absent from most of our patients with a previously identified mutation is supportive of the independent pathogenesis of those changes. An expansion was found in one case (Patient 37) who had a p.Ala321Val change in *TARDBP*. The pathogenesis of mutations in the C-terminal region of TARDBP is well established (Sreedharan *et al*., 2008; Kirby *et al*., 2010) and it is notable that the remainder of patients in our cohort with described changes in *TARDBP* did not carry the *C9ORF72* expansion. It is interesting that the age of onset in this patient was relatively young (37 years), perhaps consistent with a synergistic effect of both genetic changes on the disease pathogenesis. However, the disease duration in this patient was relatively long at 58 months. Since mutations in *TARDBP* and expansions of *C9ORF72* are both associated with neuronal cytoplasmic TDP-43 inclusions, it is conceivable that they may be affecting the same pathway in a non-additive manner. Unfortunately, no post-mortem material was available from this case.

An unreported familial ALS case with a p.Glu322Lys *OPTN* substitution also showed a *C9ORF72* expansion. While this substitution has been reported as a polymorphism in sub-Saharan African samples (Liu *et al*., 2008), it was absent from 375 neurologically normal Caucasian controls from the same geographical region. This patient had bulbar onset ALS, with onset of disease aged 50 years and a disease course of 29 months.

Dementia

An association between ALS and FTD is well established (Phukan *et al*., 2007). However, a previous population study reported that the incidence of dementia (5%) in relatives of patients with ALS was only slightly higher than in controls (Huisman *et al*., 2011) and only in first-degree relatives. ALS is thought to be the common end-point of various disease mechanisms (Ferraiuolo *et al*., 2011) and it is possible that only some of these mechanisms result in an associated dementia. Thus, disease heterogeneity may have meant that a stronger association in some patients within this population study was masked by a lack of association in other patients. In our cohort of patients with an expansion in *C9ORF72*, 35% either had a diagnosis of dementia or a family history of dementia. We propose that the subtype of ALS caused by hexanucleotide repeat expansions of *C9ORF72* shows a striking association with dementia clinically. This proposal is strongly supported by our finding of extra-motor neuropathology in all *C9ORF72* expansion cases examined at autopsy, in combination with a relative paucity of extra-motor pathology in cases found to have normal repeat lengths of *C9ORF72*. This hypothesis would be supported further by routine cognitive testing of ALS cases during life, with subsequent correlation with post-mortem neuropathology. The TDP-43 proteinopathies ALS, ALS with FTLD and pure FTLD-TDP have often been seen as a continuum of disease at both the clinical (Lillo and Hodges, 2009) and pathological levels (Mackenzie and Feldman, 2005; Geser *et al*., 2009). This has raised the question of what governs where on this spectrum

a patient with a TDP-43 proteinopathy will manifest disease pathologically, and how they will present clinically. Our data suggest that repeat expansions of *C9ORF72* have a strong influence on the clinical phenotype and the spectrum of pathology. The precise mechanism(s) by which the repeat expansion has this effect needs to be elucidated, and it will clearly be of interest to compare the distribution of pathology in brains of FTLD-TDP cases with and without the repeat expansion.

Other neurological disease

Several of the patients carrying the hexanucleotide expansion in *C9ORF72* were either diagnosed with or had a family history of other neurological, particularly neurodegenerative disease. This is illustrated by the case with both ALS and multiple sclerosis confirmed at autopsy. Previous work has suggested that the clinical spectrum of ALS may be wider than initially recognized. In particular, mutations in the *VCP* gene appear to cause a clinical spectrum including inclusion body myositis, FTD and Paget's disease (Johnson *et al*., 2010). ALS has also been reported to coexist with glaucoma in a patient with a mutation in the *OPTN* gene, although the authors went on to suggest that this may have been a coincidence (Maruyama *et al*., 2010). Analysis of our cohort suggests that expansion of *C9ORF72* may also produce a broad clinical phenotype although we have not yet demonstrated segregation of the expansion with non-ALS neurological disease.

A previous population study on the Island of Guam described coexistence of Parkinson's disease and ALS both within individuals and within families (Yanagihara *et al*., 1983). However, the results of other population studies investigating the coexistence of more than one neurodegenerative disease have been inconsistent. A recent population-based study suggested that the rate of Parkinson's disease among relatives of patients with ALS was not significantly different from controls (Huisman *et al*., 2011). However, it is noteworthy that among our patients with the hexanucleotide repeat sequence, the frequency of Parkinson's disease in relatives (6.5%) is much higher than in either the disease or control population from the Huisman study (0.9%). One of our patients with the *C9ORF72* expansion was diagnosed with both Parkinson's disease and ALS. His Parkinson's disease developed at a relatively early age of 38 years but was otherwise clinically typical, showing a good response to L-DOPA therapy; he had a deep brain stimulator inserted aged 47 years. Post-mortem he had neuropathological features of both Parkinson's disease and ALS including cell loss from the substantia nigra and 6/6 Braak grade α-syncluclein pathology with substantial staining both in the substantia nigra and in neocortical regions.

The coexistence of multiple sclerosis and ALS has been previously reported (Hader *et al*., 2008; Ismail *et al*., manuscript under review) and an increased incidence of multiple sclerosis in the offspring of patients with ALS has also been reported (Hemminki *et al*., 2009). The association of patients with ALS with the *C9ORF72* expansion and multiple sclerosis in our cohort is consistent with these reports and raises the possibility that the association between ALS and multiple sclerosis may be related to this genotype. The diagnosis of Huntington's disease in the father of a patient carrying the *C9ORF72* expansion is especially interesting given that Huntington's disease is also related to an aberrant repeat sequence. Perhaps a common underlying mechanism has led to the *de novo* occurrence of nucleotide repeat sequences in both of these individuals.

Neuropathological features

Qualitatively, the neuropathological features associated with hexanucleotide repeat expansion of *C9ORF72* are those of classical ALS with ubiquitylated, TDP-43 and OPTN positive neuronal and glial cytoplasmic inclusions in upper and lower motor neurons as well as glia. This is seen in combination with Bunina bodies and degeneration of the pyramidal tracts. What is striking is the association of pyramidal cell pathology with the expansion in the hippocampus and frontal neocortex, while very little such pathology is seen in ALS cases without the repeat expansion. This differentiation is less marked in the context of the granule cells of the hippocampal dentate gyrus, a structure that is generally believed to be one of the most sensitive extra-motor structures to TDP-43 proteinopathy (Mackenzie and Feldman, 2003; Takeda *et al*., 2009). The extra-motor pathology seen is best characterized as type B according to the scheme of Mackenzie *et al*. (2011).

The subdivision of ALS cases into two subgroups on the basis of neuropathological involvement of extra-motor (principally hippocampal and frontotemporal neocortex) has been described elsewhere (Nishihira *et al*., 2008). Data presented in this report indicate that repeat expansions in *C9ORF72* represent a major molecular basis for this pathological dichotomy.

It appears from our data derived from a large pathological cohort of ALS autopsies that the finding of extra-motor neocortical pathology, in particular ubiquitylated neuronal cytoplasmic inclusions in the CA4 subfield of the hippocampus is a relatively reliable indicator of the presence of a hexanucleotide repeat expansion in *C9ORF72*. Given that these changes can be detected at autopsy with relative ease, this neuropathological feature may be used to guide genetic investigation for the repeat expansion and thereby inform genetic counselling and further research. Interestingly, along with the characteristic hyaline conglomerate inclusions found in some patients with *SOD1* mutations (Ince *et al*., 1998), this report adds another important element to neuropathological predictors of genotype in ALS. It is also noteworthy that in the hippocampus, the majority of the neuronal cytoplasmic inclusion pathology that is evident on p62 immunohistochemistry is not apparent on TDP-43 immunohistochemistry. In fact, for the CA4 subregion (where pathology is most specific for the repeat expansion), OPTN is a more sensitive marker than TDP-43, although neither is as sensitive as p62. This raises the significant issue of what protein forms the ubiquitylated lesion in the majority of these neurons if it is not TDP-43.

Immunohistochemistry to C9ORF72 protein reveals multiple minute, puncta of labelling throughout the neuropil of grey matter structures having the appearance of synaptic labelling in neuronal processes. In this hippocampus, this is most marked in the regions that show pathology that is specific for the hexanucleotide repeat expansion. There is additionally labelling of glial cells, many of which appear to be oligodendrocytes, and around axons, corresponding to some regions of myelin. These constitute

initial observations, and the precise nature of C9ORF72 expression requires more detailed characterization, including finer localization studies, in order to begin to understand the function of the C9ORF72 protein in the CNS.

Potential mechanism of neurodegeneration in C9ORF72-related ALS

ALS and FTD now join a growing number of neurodegenerative disorders caused by expansions in repeat regions. In diseases such as Huntington's disease where the expansion is located in a coding region and is translated, the mechanism leading to neurodegeneration seems to be clear; toxic accumulation of the mutant protein that subsequently disrupts a variety of cellular processes. However, in other diseases, such as myotonic dystrophy type 1 and several of the spinocerebellar ataxias, as in *C9ORF72*, the pathological expansion is located in a non-coding region of the gene, suggesting that the RNA species itself may be toxic. Several lines of evidence are developing as to how RNA toxicity may be mediated (Todd and Paulson, 2009). In myotonic dystrophy type 1, the mutant RNA has been shown to sequester several splicing factors. This results in the formation of RNA foci within the cell nucleus and also leads to downstream messenger RNA splicing defects that are responsible for the features of muscle fibre atrophy and insulin resistance found in these patients (Fugier *et al*., 2011). Consistent with this as a potential mechanism occurring in ALS/FTD cases with expansions in *C9ORF72*, DeJesus-Hernandez *et al*. (2011) demonstrated the presence of increased numbers of neuronal RNA foci in ALS cases with the expansion. Additionally, levels of C9ORF72 protein in lymphoblastoid cell lines and frontal cortex samples taken from patients with FTD with pathological expansions were not significantly different to those from patients with FTD without expansions (DeJesus-Hernandez *et al*., 2011). The splicing factor sequestration hypothesis also links *C9ORF72* expansions to the TDP-43 pathology present in these cases; the latter being implicated in alternative splicing of multiple transcripts (Buratti and Baralle, 2001; Bose *et al*., 2008). However, to date little is known about the function of the C9ORF72 protein and it remains possible that defective splicing of its own messenger RNA or transcriptional silencing resulting in haploinsufficiency, might contribute to the neurodegenerative process.

Conclusion

This is the first detailed report of the clinical and pathological phenotypes in a large cohort of 62 ALS/MND cases with pathological hexanucleotide expansions in the *C9ORF72* gene. This genetic variant is common, accounting for 43% of familial ALS and 7% of sporadic ALS cases in this Caucasian cohort from Northern England. Clinical features of note include the prevalence of dementia in the ALS cases and close family members, and a younger age of onset and more rapid disease progression compared with other ALS/MND subtypes. Pathologically, these cases showed classical ALS changes, with TDP-43 positive inclusions. Extra-motor system pathology was strikingly over-represented in the C9*ORF*72 positive cases, suggesting that this genotype may be one predictor of an individual ALS patient's location along the anatomical spectrum of ALS pathology. Neuronal cytoplasmic inclusions within the CA4 hippocampal region may be sufficiently characteristic to allow prediction of mutation status.

Acknowledgements

We thank all the ALS/MND patients from the Sheffield and Newcastle Care and Research Centres for Motor Neurone Disorders and their family members who contributed so generously to research by the donation of biosamples.

Funding

European Community's Health 7th Framework Programme (FP 7 2007–2013 under grant agreement no. 259867 to P.J.S. and J.K.). MND Association (Shaw/Nov02/6700/3 to P.J.S.); Wellcome Trust (WT070122MF to P.J.S.); MND Association/Medical Research Council Lady Edith Wolfson Fellowship award (G0 800380 to J.R.H.); Wellcome Trust/MRC Strategic Neurodegeneration Award (WT 089698/Z/09/Z to J.H.) and Intramural Research Programs of the NIH, National Institute on Aging (Z01-AG000949-02 to B.J.T.) and National Institute for Neurological Disorders and Stroke (to B.J.T.).

Supplementary material

Supplementary material is available at *Brain* online.

References

Bose JK, Wang IF, Hung L, Tarn WY, Shen CK. TDP-43 overexpression enhances exon 7 inclusion during the survival of motor neuron pre-mRNA splicing. J Biol Chem 2008; 283: 28852–9.

Brooks BR, Miller RG, Swash M, Munsat TL. World Federation of Neurology Research Group on Motor Neuron Diseases. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotroph Lateral Scler Other Motor Neuron Disord 2000; 1: 293–9.

Buratti E, Baralle FE. Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. J Biol Chem 2001; 276: 36337–43.

DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron 2011. Advance Access published on October 27, 2011, doi:10.1016/j.neuron.2011.09.011.

Elden AC, Kim HJ, Hart MP, Chen-Plotkin AS, Johnson BS, Fang X, et al. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. Nature 2010; 466: 1069–75.

Fecto F, Siddique T. Making connections: Pathology and genetics link amyotrophic lateral sclerosis with frontotemporal lobe dementia. J Mol Neurosci 2011; 45: 663–75.

Ferraiuolo L, Kirby J, Grierson AJ, Sendtner M, Shaw PJ. Molecular and cellular pathways of motor neuron injury in amyotrophic lateral sclerosis. Nat Rev Neurol 2011; 7: 616–30.

Fugier C, Klein AF, Hammer C, Vassilopoulos S, Ivarsson Y, Toussaint A, et al. Misregulated alternative splicing of BIN1 is associated with T

tubule alterations and muscle weakness in myotonic dystrophy. Nat Med 2011; 17: 720–5.
Geser F, Martinez-Lage M, Robinson J, Uryu K, Neumann M, Brandmeir MJ, et al. Clinical and pathological continuum of multisystem TDP-43 proteinopathies. Arch Neurol 2009; 66: 180–9.
Hader WJ, Rpzdilsky B, Nair CP. The concurrence of multiple sclerosis and amyotrophic lateral sclerosis. Can J Neurol Sci 1986; 13: 66–9.
Hemminki K, Li X, Sundquist J, Sundquist K. Familial risks for amyotrophic lateral sclerosis and autoimmune diseases. Neurogenetics 2009; 10: 111–6.
Hewitt C, Kirby J, Highley JR, Hartley JR, Hibberd R, Hollinger HC, et al. Novel FUS/TLS mutations and pathology in familial and sporadic amyotrophic lateral sclerosis. Arch Neurol 2010; 67: 455–61.
Hosler BA, Siddique T, Sapp PC, Sailor W, Huang MC, Hossain A, et al. Linkage of familial amyotrophic lateral sclerosis with frontotemporal dementia to chromosome 9q21–q22. JAMA 2000; 284: 1664–9.
Huisman M, de Jong S, Verwijs M, Schelhaas H, van der Kooi A, de Visser M, et al. Family history of neurodegenerative and vascular diseases in ALS: A population-based study. Neurology 2011; 77: 1363–9.
Ince PG, Tomkins J, Slade JY, Thatcher NM, Shaw PJ. Amyotrophic lateral sclerosis associated with genetic abnormalities in the gene encoding Cu/Zn superoxide dismutase: Molecular pathology of five new cases, and comparison with previous reports and 73 sporadic cases of ALS. J Neuropathol Exp Neurol 1998; 57: 895–904.
Johnson JO, Mandrioli J, Benatar M, Abramzon Y, Van Deerlin VM, Trojanowski J, et al. Exome sequencing reveals VCP mutations as a cause of familial ALS. Neuron 2010; 68: 857–64.
Kirby J, Goodall E, Smith W, Highley JR, Masanzu R, Hartley J, et al. Broad clinical phenotypes associated with TAR-DNA binding protein mutations in amyotrophic lateral sclerosis. Neurogenetics 2010; 11: 217–25.
Laaksovirta H, Peuralinna T, Schymick T, Scholz SW, Lai SL, Myllykangas L, et al. Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. Lancet Neurol 2010; 9: 978–85.
Lambrechts D, Storkebaum E, Morimoto M, Del-Favero J, Desmet F, Marklund SL, et al. VEGF is a modifier of amyotrophic lateral sclerosis in mice and humans and protects motoneurons against ischemic death. Nat Genet 2003; 34: 383–94.
Lillo P, Hodges JR. Frontotemporal dementia and motor neurone disease: Overlapping clinic-pathological disorders. J Clin Neurosci 2009; 16: 1131–5.
Liu Y, Akafo S, Santiago-Turla C, Cohen CS, Laracque-Abramson KR, Qin X, et al. Optineurin coding variants in Ghanaian patients with primary open-angle glaucoma. Mol Vis 2008; 14: 2367–72.
Mackenzie IRA, Feldman H. Neuronal intranuclear inclusions distinguish familial FTD-MND type from sporadic cases. Acta Neuropathol 2003; 105: 543–8.
Mackenzie IRA, Feldman HH. Ubiquitin immunohistochemistry suggests classic motor neuron disease, motor neuron disease with dementia, and frontotemporal dementia of the motor neuron disease type represent a clinicopathologic spectrum. J Neuropathol Exp Neurol 2005; 64: 730–9.
Mackenzie IRA, Neumann M, Baborie A, Sampathu DM, Du Plessis D, Jaros E, et al. A harmonized classification system for FTLD-TDP pathology. Acta Neuropathol 2011; 122: 111–3.
Maruyama H, Morino H, Ito H, Izumi Y, Kato H, Watanabe Y, et al. Mutations of optineurin in amyotrophic lateral sclerosis. Nature 2010; 465: 223–6.
Mok K, Traynor BJ, Schymick J, Tienari PJ, Laaksovirta H, Peuralinna T, et al. The chromosome 9 ALS and FTD locus is probably derived from a single founder. Neurobiol Aging 2011. Advance Access published on September 16, 2011, doi:10.1016/j.neurobiolaging.2011.08.05.
Nishihira Y, Tan CF, Onodera O, Toyoshima Y, Yamada M, Morita T, et al. Sporadic amyotrophic lateral sclerosis: two pathological patterns shown by analysis of distribution of TDP-43-immunoreactive neuronal and glial cytoplasmic inclusions. Acta Neuropathol 2008; 116: 169–82.
Phukan J, Elamin M, Bede P, Jordan N, Gallagher L, Byrne S, et al. The syndrome of cognitive impairment in amyotrophic lateral sclerosis: a population-based study. J Neurol Neurosurg Psychiatry 2011. Advance Access published on August 11, 2011, doi:10.1136/jnnp-2011-300188.
Phukan J, Pender NP, Hardiman O. Cognitive impairment in amyotrophic lateral sclerosis. Lancet Neurol 2007; 6: 994–1003.
Renton AE, Majounie E, Waite A, Simon-Sanchez J, Rollinson S, Gibbs JR, et al. A Hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-Linked ALS-FTD. Neuron 2011. Advance Access published on September 21, 2011, doi:10.1016/j.neuron.2011.09.010.
Sreedharan J, Blair IP, Tripathi VB, Hu X, Vance C, Rogelj B, et al. TDP-43 Mutations in familial and sporadic amyotrophic lateral sclerosis. Science 2008; 319: 1668–72.
Takeda T, Uchihara T, Arai N, Mizutani T, Iwata M. Progression of hippocampal degeneration in amyotrophic lateral sclerosis with or without memory impairment: distinction from Alzheimer disease. Acta Neuropathol 2009; 117: 35–44.
Ticozzi N, Tiloca C, Morelli C, Colombrita C, Poletti B, Doretti A, et al. Genetics of familial amyotrophic lateral sclerosis. Arch Ital Biol 2011; 149: 65–82.
Todd PK, Paulson HL. RNA-mediated neurodegeneration in repeat expansion disorders. Ann Neurol 2009; 67: 291–300.
van Es MA, Veldink JH, Saris CG, Blauw HM, van Vught PW, Birve A, et al. Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. Nat Genet 2009; 41: 1083–7.
Yanagihara R, Garruto R, Gajdusek D. Epidemiological surveillance of amyotrophic lateral sclerosis and parkinsonism-dementia in the Commonwealth of the Northern Mariana Islands. Ann Neurol 1983; 13: 79–86.

Frequency of the *C9orf72* hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study

Elisa Majounie, Alan E Renton*, Kin Mok*, Elise G P Dopper*, Adrian Waite*, Sara Rollinson*, Adriano Chiò*, Gabriella Restagno*, Nayia Nicolaou*, Javier Simon-Sanchez*, John C van Swieten*, Yevgeniya Abramzon, Janel O Johnson, Michael Sendtner, Roger Pamphlett, Richard W Orrell, Simon Mead, Katie C Sidle, Henry Houlden, Jonathan D Rohrer, Karen E Morrison, Hardev Pall, Kevin Talbot, Olaf Ansorge, The Chromosome 9-ALS/FTD Consortium†, The French research network on FTLD/FTLD/ALS†, The ITALSGEN Consortium†, Dena G Hernandez, Sampath Arepalli, Mario Sabatelli, Gabriele Mora, Massimo Corbo, Fabio Giannini, Andrea Calvo, Elisabet Englund, Giuseppe Borghero, Gian Luca Floris, Anne M Remes, Hannu Laaksovirta, Leo McCluskey, John Q Trojanowski, Vivianna M Van Deerlin, Gerard D Schellenberg, Michael A Nalls, Vivian E Drory, Chin-Song Lu, Tu-Hsueh Yeh, Hiroyuki Ishiura, Yuji Takahashi, Shoji Tsuji, Isabelle Le Ber, Alexis Brice, Carsten Drepper, Nigel Williams, Janine Kirby, Pamela Shaw, John Hardy, Pentti J Tienari*, Peter Heutink*, Huw R Morris*, Stuart Pickering-Brown*, Bryan J Traynor**

Summary

Background We aimed to accurately estimate the frequency of a hexanucleotide repeat expansion in *C9orf72* that has been associated with a large proportion of cases of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD).

Methods We screened 4448 patients diagnosed with ALS (El Escorial criteria) and 1425 patients with FTD (Lund-Manchester criteria) from 17 regions worldwide for the GGGGCC hexanucleotide expansion using a repeat-primed PCR assay. We assessed familial disease status on the basis of self-reported family history of similar neurodegenerative diseases at the time of sample collection. We compared haplotype data for 262 patients carrying the expansion with the known Finnish founder risk haplotype across the chromosomal locus. We calculated age-related penetrance using the Kaplan-Meier method with data for 603 individuals with the expansion.

Findings In patients with sporadic ALS, we identified the repeat expansion in 236 (7·0%) of 3377 white individuals from the USA, Europe, and Australia, two (4·1%) of 49 black individuals from the USA, and six (8·3%) of 72 Hispanic individuals from the USA. The mutation was present in 217 (39·3%) of 552 white individuals with familial ALS from Europe and the USA. 59 (6·0%) of 981 white Europeans with sporadic FTD had the mutation, as did 99 (24·8%) of 400 white Europeans with familial FTD. Data for other ethnic groups were sparse, but we identified one Asian patient with familial ALS (from 20 assessed) and two with familial FTD (from three assessed) who carried the mutation. The mutation was not carried by the three Native Americans or 360 patients from Asia or the Pacific Islands with sporadic ALS who were tested, or by 41 Asian patients with sporadic FTD. All patients with the repeat expansion had (partly or fully) the founder haplotype, suggesting a one-off expansion occurring about 1500 years ago. The pathogenic expansion was non-penetrant in individuals younger than 35 years, 50% penetrant by 58 years, and almost fully penetrant by 80 years.

Interpretation A common Mendelian genetic lesion in *C9orf72* is implicated in many cases of sporadic and familial ALS and FTD. Testing for this pathogenic expansion should be considered in the management and genetic counselling of patients with these fatal neurodegenerative diseases.

Funding Full funding sources listed at end of paper (see Acknowledgments).

Lancet Neurol 2012; 11: 323–30

Published Online March 9, 2012 DOI:10.1016/S1474-4422(12)70043-1

See Comment page 297

*Authors contributed equally

†Members listed in the appendix

Molecular Genetics Unit (E Majounie PhD, D G Hernandez MSc, S Arepalli MS, M A Nalls PhD), **Neuromuscular Diseases Research Unit** (A E Renton PhD, Y Abramzon, J O Johnson PhD, B J Traynor MD), **Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA; Department of Molecular Neuroscience and Reta Lila Weston Laboratories** (K Mok MSc, K C Sidle MD, Prof J Hardy PhD), **Department of Clinical Neurosciences** (R W Orrell MD), **MRC Prion Unit, Department of Neurodegenerative Disease** (S Mead MD), **Department of Molecular Neurosciences and MRC Centre for Neuromuscular Diseases** (Prof H Houlden MD), **and Department of Neurodegenerative Disease, Dementia Research Centre** (J D Rohrer MD), **Institute of Neurology, University College London, Queen Square House, London, UK; Department of Clinical Genetics, Section of Medical Genomics, and Alzheimer Center, VU University Medical Centre, Amsterdam, Netherlands** (E G P Dopper, N Nicolaou MSc, J Simon-Sanchez PhD, Prof J C van Swieten MD, Prof P Heutink PhD); **Department of Neurology,**

Introduction

Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disease characterised by rapidly progressive paralysis and death from respiratory failure, typically within 3 years of symptom onset. The disease is inherited in about 5% of cases, following a clear Mendelian pattern, whereas most cases are classified as sporadic because they seem to arise at random.[1] Substantial progress has been made in understanding the genetic underpinnings of familial ALS.[2] By contrast, the causes of sporadic or idiopathic ALS are far less well understood. Mutations in the known familial ALS genes—*SOD1*, *FUS*, and *TDP-43*—occur only rarely in sporadic cases (each accounting for less than 1·0% of cases);[3–5] genome-wide association studies have identified few risk loci, and these have proved difficult to replicate.[6]

Frontotemporal dementia (FTD) is a degenerative disorder of the frontal and anterior temporal lobes, and is a common form of dementia affecting individuals younger than 65 years. The syndrome is characterised clinically by initial behavioural disturbances, followed by cognitive decline leading to dementia and death within a median of 7 years from symptom onset. Akin to ALS and other neurodegenerative diseases, a large proportion (~60·0%) of these cases are categorised as sporadic, and the causes of this idiopathic form of disease are largely unknown.[7] A growing consensus

Erasmus MC–University Medical Center Rotterdam, Rotterdam, Netherlands (E G P Dopper, N Nicolaou, J Simon-Sanchez, J C van Swieten); **MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Cardiff, UK** (A Waite PhD, N Williams PhD, H R Morris MD); **Faculty of Human and Medical Sciences, University of Manchester, Manchester, UK** (S Rollinson PhD, Prof S Pickering-Brown PhD); **Department of Neuroscience, University of Turin, Turin, Italy** (A Chiò MD, A Calvo MD); **Molecular Genetics Unit, Department of Clinical Pathology, Azienda Ospedaliera Ospedale Infantile Regina Margherita Sant Anna, Turin, Italy** (G Restagno MD); **Institute for Clinical Neurobiology, University of Würzburg, Würzburg, Germany** (Prof M Sendtner MD, C Drepper PhD); **Department of Pathology, Sydney Medical School, The University of Sydney, NSW, Australia** (R Pamphlett MD); **Department of Neurology, Institute of Biomedical Research, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK** (Prof K E Morrison MD); **Neurology–University Hospitals Birmingham NHS Foundation Trust, Queen Elizabeth Hospital, Queen Elizabeth Medical Centre, Birmingham, UK** (H Pall MD); **Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, University of Oxford, Oxford, UK** (Prof K Talbot MD, O Ansorge MD); **Neurological Institute, Catholic University and ICOMM Association for ALS Research, Rome, Italy** (M Sabatelli MD); **ALS Center, Salvatore Maugeri Foundation, Milan, Italy** (G Mora MD); **NeuroMuscular Omnicentre, Niguarda Ca' Granda Hospital, Milan, Italy** (M Corbo MD); **Department of Neurological, Neurosurgical and Behavioural Sciences, Neurology Section, University of Siena, Siena, Italy** (F Giannini MD); **Department of Pathology, Lund University, Regional Laboratories Region Skåne, Lund, Sweden** (E Englund MD); **Department**

suggests that ALS and FTD form part of a continuum of neurological diseases that share a common pathological background, consisting of TAR DNA-binding protein 43 (TDP-43)-positive inclusions within the CNS.[8]

We recently reported that a large hexanucleotide repeat expansion located within the non-coding portion of *C9orf72* is the cause of chromosome 9-linked ALS and FTD.[9,10] This genetic lesion accounted for a large proportion (~40·0%) of familial cases of ALS and FTD. The same mutation was present in nearly a quarter of apparently sporadic cases of ALS and FTD in the genetically homogeneous Finnish population, and in 4·1% of sporadic cases of ALS and 3·0% cases of sporadic FTD from the USA. However, these estimates were based on relatively small cohorts drawn from a small number of institutions.

These findings prompted us to aim to estimate the frequency of this *C9orf72* hexanucleotide repeat expansion more accurately, in a large cohort of European and US patients with sporadic ALS and sporadic FTD. We also examined the occurrence of this mutation in diverse non-white populations around the world.

Methods

Participants and study design

In this cross-sectional study, we screened 4448 patients diagnosed with ALS and 1425 patients diagnosed with FTD from 17 distinct regions worldwide. The appendix shows ethnic origin and clinical features of the patients. 3860 patients had sporadic ALS, 1022 had sporadic FTD, 588 had familial ALS, and 403 had familial FTD. Data for 401 Finnish patients with ALS, 233 other Europeans with familial ALS, 75 Finnish patients with FTD, 340 Dutch patients with FTD, and 420 English patients with FTD have been published previously.[10–12] All these cohorts were analysed to provide a comprehensive assessment of the global frequency of the expansion.

Patients with ALS were diagnosed according to the El Escorial criteria,[13] and patients with FTD were diagnosed according to the Lund-Manchester criteria.[14] We classified patients' disease as familial in nature on the basis of a diagnosis of ALS or FTD in any other family member (irrespective of relationship), as reported at the time of sample collection. We based ethnic and racial classification on self-reports from patients at the time of sample collection. Case numbers

	Sporadic ALS			Sporadic FTD		
	n	Carriers	% (95% CI)	n	Carriers	% (95% CI)
Europe*						
Finnish	289	61	21·1% (16·5–26·3)	48	9	18·8% (8·9–32·6)
Swedish	..	..	..	6	0	0% (0·0–45·9)
English	916	62	6·8% (5·2–8·6)	543	31	5·7% (3·9–8·0)
German	421	22	5·2% (3·3–7·8)	..	..	..
Dutch	..	..	..	224	5	2·2% (0·7–5·1)
French	..	..	..	150	14	9·3% (5·2–15·2)
Italian	465	19	4·1% (2·5–6·3)	..	..	..
Sardinian	129	10	7·8% (3·8–13·8)	10	0	0% (0·0–30·8)
Moldovan	3	0	0% (0·0–70·8)	..	..	..
Total (Europe)	2223	174	7·8% (6·7–9·0)	981	59	6·0% (4·6–7·7)
USA						
White	890	48	5·4% (4·0–7·1)	..	..	..
Hispanic	72	6	8·3% (3·1–17·3)	..	..	..
Black	49	2	4·1% (0·5–14·0)	..	..	..
Native American	3	0	0% (0·0–70·8)	..	..	..
Total (USA)	1014	56	5·5% (4·2–7·1)	..	..	..
Rest of the world						
Middle Eastern*	1	0	0% (0·0–97·5)	..	..	..
Indian	31	0	0% (0·0–11·2)	31	0	0% (0·0–11·2)
Asian	238	0	0% (0·0–1·5)	10	0	0% (0·0–30·8)
Pacific Islander/Guam	90	0	0% (0·0–4·0)	..	..	..
Australian*	263	14	5·3% (2·9–8·8)	..	..	..
Overall	3860	244	6·3% (5·6–7·1)	1022	59	5·8% (4·4–7·4)

Data for Finnish (289 with ALS and 48 with FTD), English (333 with FTD), and Dutch (224 with FTD) patients were previously published,[10–12] but are included here to establish global frequencies. ALS=amyotrophic lateral sclerosis. FTD=frontotemporal dementia. *All self-reported as white.

Table 1: **Frequency of the pathogenic GGGGCC hexanucleotide repeat expansion of *C9orf72* in patients diagnosed with sporadic ALS or sporadic FTD classified by region**

	Familial ALS			Familial FTD		
	n	Carriers	% (95% CI)	n	Carriers	% (95% CI)
Europe*						
Finnish	112	52	46·4% (37·0–56·1)	27	13	48·1% (28·7–68·0)
Swedish	..	..	..	1	1	100·0% (2·5–100·0)
English	98	45	45·9% (35·8–56·3)	170	28	16·5% (11·2–22·9)
Irish	1	1	100·0% (2·5–100·0)	..	..	
German	69	15	21·7% (12·7–33·3)	29	4	13·8% (3·9–31·7)
Dutch	..	..	..	116	30	25·9% (18·2–34·8)
French	..	..	..	50	22	44·0% (30·0–58·7)
Italian	90	34	37·8% (27·8–48·6)	..	..	..
Sardinian	19	11	57·9% (33·5–79·7)	7	1	14·3% (0·4–57·9)
Total (Europe)	389	158	40·6% (35·7–45·7)	400	99	24·8% (20·6–29·3)
USA*	163	59	36·2% (28·8–44·1)	..	..	..
Rest of the world						
Middle Eastern*	2	0	0% (0·0–84·2)	..	..	..
Israeli*	14	3	21·4% (4·7–50·8)	..	..	..
Asian	20	1	5·0% (0·1–24·9)	3	2	66·7% (9·4–99·2)
Overall	588	221	37·6% (33·7–41·6)	403	101	25·1% (20·9–29·6)

Data for Finnish (112 with ALS and 27 with FTD), English (87 with FTD), German (41 with ALS), Italian (29 with ALS), US (163 with ALS), and Dutch (116 with FTD) patients were previously published,[10–12] but are included here to establish global frequencies. ALS=amyotrophic lateral sclerosis. FTD=frontotemporal dementia. *All self-reported as white.

Table 2: **Frequency of the pathogenic GGGGCC hexanucleotide repeat expansion of *C9orf72* in patients diagnosed with familial ALS and familial FTD classified by region**

listed for European countries and Australia and the Middle East refer to self-reported white individuals from that region. Italian data are from a population-based cohort that had been collected through the Piemonte ALS Registry, an ongoing population-based epidemiological study of ALS based in northwestern Italy.[15] The remaining cohorts were recruited through medical centres and from repositories in various countries.

We also screened 2585 neurologically healthy control individuals from Australia (213 patients), Finland (478), Germany (309), the Human Gene Diversity Panel (300), mainland Italy (354), Sardinia (87), and the USA (844) for presence of the pathogenic repeat expansion. 1167 of these individuals have been reported elsewhere.[10] None of the control individuals had been diagnosed with ALS, FTD, dementia, or any other neurodegenerative disease. Ethics committees from the respective institutions approved the study, and written informed consent was obtained from all patients and control individuals.

Procedures

We used our previously described[10] repeat-primed PCR assay to screen patients and control individuals for the presence of the chromosome 9p21 GGGGCC hexanucleotide repeat expansion (see appendix for technical details). The assay allows samples to be categorised into those that carry a pathogenic repeat expansion (>30 repeats) and those that carry only wild-type alleles (<20 repeats).

For haplotype analysis, we analysed genome-wide single-nucleotide polymorphism (SNP) data from 262 patients who carried the repeat expansion. We previously reported the identification in the Finnish population of a 42-SNP founder haplotype across the 232 kb block of chromosome 9p21 where the pathogenic hexanucleotide expansion was ultimately established.[16,17] In this study, we used a custom perl software script to compare unphased sample genotype data with the 42-SNP founder risk haplotype.[16]

We estimated mutation ages for all populations separately with the DMLE+ version 2.3 Bayesian linkage disequilibrium gene mapping package.[18] Mutation ages were iterated for 10 000 burn-in iterations and a further 10 000 iterations of the maximum-likelihood model. To obtain generalisable estimates of age of the repeat per population, we used median values of binned estimates passing the α threshold of 0·05 per iteration.

Statistical analysis

We calculated 95% CIs for proportions with the Clopper-Pearson exact method. We estimated penetrance of the GGGGCC hexanucleotide repeat expansion in relation to the patients' age on the basis of data available for 603 mutant-gene carriers with the Kaplan-Meier method using the survival package within R statistical software (version 2.9.0), but substituting patient age at symptom onset for survival time.[19] We assessed differences between groups with the χ^2 test for discrete variables such as sex, family history, and site of onset.

of Neurology, Azienda Universitaria-Ospedaliera di Cagliari and University of Cagliari, Cagliari, Italy (G Borghero MD, G L Floris MD); **Institute of Clinical Medicine, Neurology, University of Oulu and Clinical Research Center, Oulu University Hospital, Oulu, Finland** (Prof A M Remes MD); **Department of Neurology, Helsinki University Central Hospital and Molecular Neurology Programme, Biomedicum, University of Helsinki, Helsinki, Finland** (H Laaksovirta MD, P J Tienari MD); **Department of Neurology** (L McCluskey MD), **Department of Pathology and Laboratory Medicine** (Prof J Q Trojanowski MD, V M Van Deerlin MD, Prof G D Schellenberg PhD), **University of Pennsylvania, Philadelphia, PA, USA; Department of Neurology, Tel-Aviv Sourasky Medical Center, Tel-Aviv, Israel** (V E Drory MD); **Department of Neurology, Chang Gung Memorial Hospital at Linkou Medical Center and Chang Gung University, Taoyuan, Taiwan** (Prof C-S Lu MD, T-H Yeh MD); **Neuroscience Research Center, Chang Gung Memorial Hospital at Linkou Medical Center, Taoyuan, Taiwan** (C-S Lu, T-H Yeh); **Department of Neurology, University of Tokyo Hospital, 7–3-1 Hongo, Bunkyo-ku, Tokyo, Japan** (H Ishiura MD, Y Takahashi MD, Prof S Tsuji MD); **Université Pierre et Marie Curie-Paris 6, Centre de Recherche de l'Institut du Cerveau et de la Moelle épinière, Paris, France** (I Le Ber MD, Prof A Brice MD); **INSERM, U975, Paris, France** (I Le Ber, A Brice); **CNRS, UMR 7225, Paris, France** (I Le Ber, A Brice); **Department of Neuroscience, University of Sheffield, Sheffield, UK** (J Kirby PhD, Prof P Shaw MD); **Neurology (C4), University Hospital of Wales, Cardiff, UK** (H R Morris MD); **Department of Neurology, Royal Gwent Hospital, Aneurin Bevan Local Health Board, Gwent, UK** (H R Morris); **and Department of Neurology, Brain Sciences Institute, Johns Hopkins Hospital, Baltimore, MD, USA** (B J Traynor)

Correspondence to:
Dr Bryan J Traynor, Neuromuscular Diseases Research Unit, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, 35 Convent Drive, Room 1A-1000, Bethesda, MD 20892, USA
traynorb@mail.nih.gov

See Online for appendix

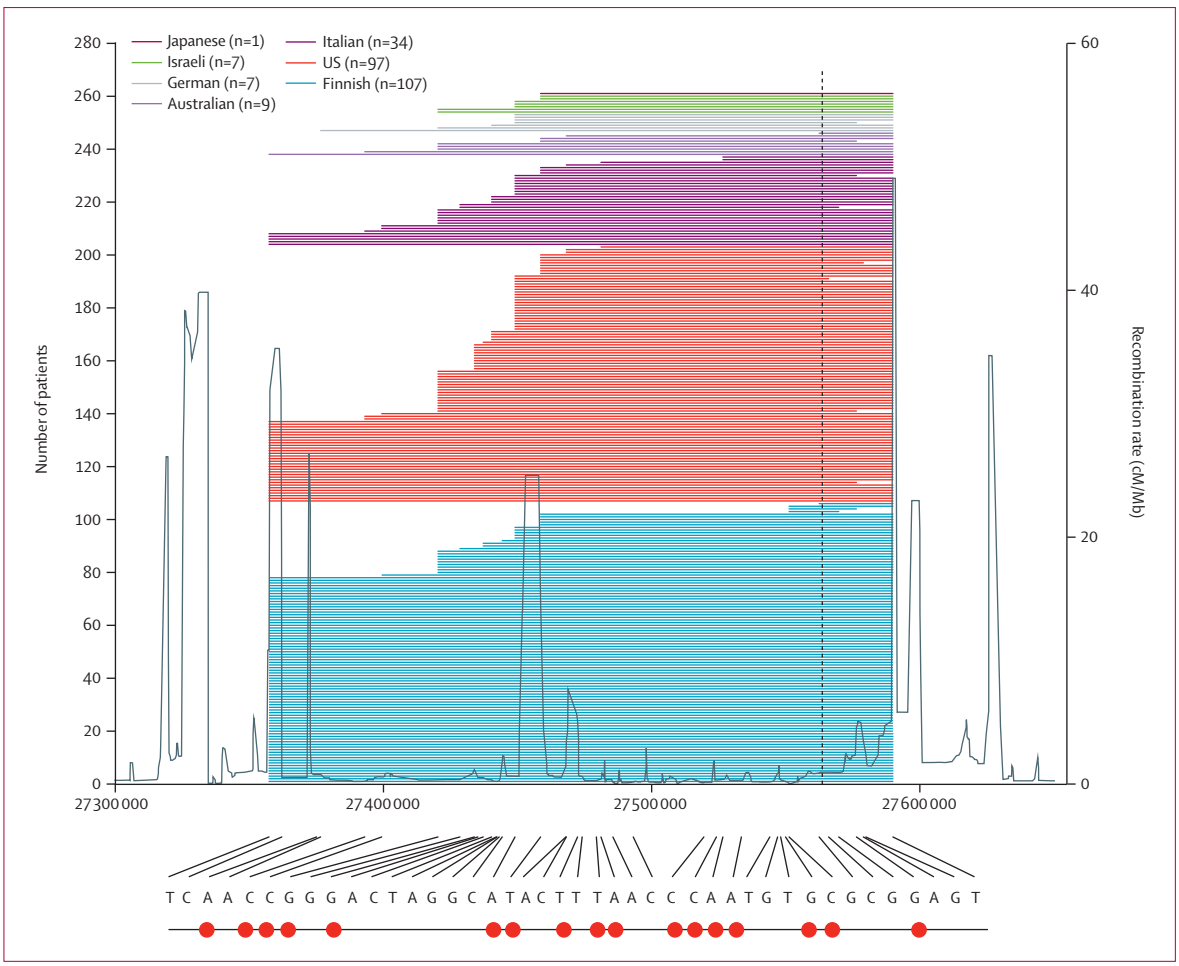


Figure 1: **Finnish risk haplotypes across the chromosome 9p21 region in 262 patients with amyotrophic lateral sclerosis and the *C9orf72* mutation**
The previously identified Finnish risk haplotype is shown below the graph (27 357 278–27 589 746 bp; NCBI build 36; 42 single nucleotide polymorphisms [SNPs]).[16] Underneath the haplotype is a binary representation of the same data, with red circles at SNP positions where the haplotype has the less common allele at that site. In the graph, individual patients are shown as horizontal lines showing the extent to which they share the risk haplotype. The vertical black dashed line shows the location of the *C9orf72* hexanucleotide repeat expansion. Recombination rates (centimorgans per megabase [cM/Mb]) from phase 2 Centre d'Etude du Polymorphisme Humain (CEPH) samples of HapMap are shown with a grey line.

Role of the funding source
The sponsors of the study had no role in study design, data collection, analysis, or interpretation, writing of the report, or in the decision to submit the paper for publication. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results
Table 1 and the appendix show the frequency of the *C9orf72* hexanucleotide repeat expansion in patients diagnosed with sporadic ALS and sporadic FTD from different geographical regions. Data for 289 patients with sporadic ALS and 605 with sporadic FTD have been reported elsewhere.[10–12] The pathogenic expansion was identified in 236 (7·0%) of 3377 white patients from the USA, Europe, the Middle East, and Australia, two (4·1%) of 49 black patients from the USA, and six (8·3%) of 72 Hispanic patients from the USA who were diagnosed with sporadic ALS. The rate of the pathogenic expansion was lower in sporadic FTD: 59 (6·0%) of 981 white patients from Europe carried the mutation. By contrast, the GGGGCC repeat expansion was not present in patients of Native American, Asian, or Pacific Islander origin who had sporadic disease (table 1), although this might reflect the smaller size of the cohorts screened in these populations.

In addition to sporadic cases, we screened 588 familial cases of ALS and 403 familial cases of FTD for the presence of the *C9orf72* repeat expansion (table 2, appendix). Of these, 345 patients with familial ALS and 230 with familial FTD have been reported elsewhere.[10–12] Overall, 221 (37·6%) of 588 patients with familial ALS and 101 (25·1%) of 403 patients with familial FTD carried the genetic lesion, reinforcing our previous findings that this mutation was responsible for an

unparalleled proportion of cases of these diseases.[10] We identified one Japanese individual diagnosed with familial ALS who carried the hexanucleotide repeat expansion. We also showed that one patient with familial FTD from Lund, Sweden, carried the expansion, suggesting that the chromosome 9p21 genetic lesion might be responsible for the geographical cluster of patients with FTD noted in that region.[20]

Of 2585 neurologically healthy control samples screened for the *C9orf72* repeat expansion, five (0·2%) were carriers: two were previously reported elderly individuals from Finland,[10] and the other three were individuals younger than 40 years from Germany and the USA (appendix).

Within Europe, the highest mutation frequency was noted in the Finnish population (21·1% of patients with sporadic ALS and 18·8% of patients with sporadic FTD).[10] About 6% of patients with sporadic ALS from Germany and England carried the expansion, whereas Italian patients with ALS had a lower rate (4·1%). 7·8% of patients with sporadic ALS from the genetically isolated island population of Sardinia had the mutation and the Dutch population had the lowest detected rate observed in European countries (2·2% of sporadic cases of FTD). White populations from Australia and the USA had an intermediate rate, with about 5·0% of patients with sporadic ALS carrying the pathogenic repeat expansion, perhaps because of the population and immigration histories of these countries.

Haplotype analysis suggested that every patient carrying the pathogenic GGGGCC repeat expansion also shared the Finnish founder risk haplotype, at least in part (figure 1). Furthermore, patients with sporadic and familial disease carried the same founder risk haplotype. These findings suggest that the pathogenic hexanucleotide repeat expansion in *C9orf72* might have occurred on one occasion in human history and subsequently disseminated throughout these populations. Analysis of haplotype sharing between these cases estimated the age of *C9orf72* repeat expansion to be about 1500 years old (representing a median of 100·5 generations [IQR 57·6–127·6], assuming a generation is 15 years old).

In analysis of age-related penetrance (figure 2), the pathogenic expansion was non-penetrant in carriers who were younger than 35 years of age, increasing to 50% penetrance by 58 years, and to almost full penetrance by 80 years. We noted no difference between disease penetrance according to familial status, ALS or FTD diagnosis, sex, or age of symptom onset in patients with ALS or FTD (appendix).

Table 3 shows clinical details of patients carrying the hexanucleotide repeat expansion. Patients with ALS and the pathogenic repeat expansion were more likely to be female (p=0·0008), have a family history of disease (p<0·0001), and to have bulbar-onset disease (p=0·0011) than were patients who did not carry the expansion. Patients with FTD carrying the repeat expansion were also more likely to have a family history of disease (p<0·0001) and to present with behavioural variant FTD (p<0·0001).

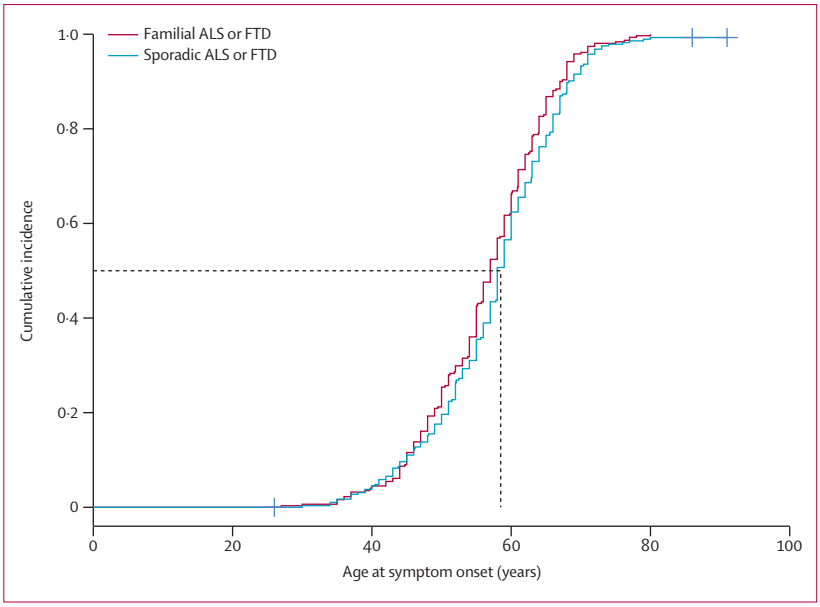


Figure 2: **Age-related penetrance of the GGGGCC hexanucleotide repeat expansion in *C9orf72***
Kaplan-Meier analysis of 603 mutant-gene carriers (212 patients with familial amyotrophic lateral sclerosis, 234 with sporadic amyotrophic lateral sclerosis, 99 with familial frontotemporal dementia, 53 with sporadic frontotemporal dementia, and five neurologically healthy controls). Age-related penetrance (ie, the proportion of mutant-gene carriers with manifestations of the disease by a given age) rose steadily, from 10% in patients younger than 45 years to almost 100% by the age of 80 years. The dotted lines shows the age at which 50% of the cohort developed symptoms. Vertical blue lines show censored events.

	Amyotrophic lateral sclerosis		Frontotemporal dementia	
	With expansion (n=465)*	Without expansion (n=3983)†	With expansion (n=160)‡	Without expansion (n=1265)§
Mean age at onset (range; SD)	56·8 (27·0–80·0; 9·1)	58·7 (4·0–93·0; 12·8)	57·5 (30·0–76·3; 8·3)	60·0 (23·0–87·0; 8·8)
Sex, male	232 (50·1%)	2251 (58·4%)	87 (54·4%)	683 (55·4%)
Positive family history	221 (47·5%)	367 (9·2%)	101 (63·1%)	302 (23·9%)
Presentation				
Bulbar	139 (33·1%)	933 (26·0%)	..	..
Limb	281 (66·9%)	2655 (74·0%)	..	..
Behavioural variant	..	..	106 (85·5%)	685 (65·6%)
Progressive non-fluent aphasia	..	..	11 (8·9%)	165 (15·8%)
Semantic dementia	..	..	7 (5·6%)	195 (18·6%)

Data are mean (range; SD) or n (%). *Data not available for age at onset for 19 patients and site of onset for 45 patients. †Data not available for age at onset for 305 patients, sex for 130 patients, and site of onset for 395 patients. ‡Data not available for age at onset for eight patients and site of onset for 36 patients. §Data not available for age at onset for 71 patients, sex for32 patients, and site of onset for 220 patients.

Table 3: **Demographic and clinical features of patients classified by diagnosis and by carrier status for the GGGGCC hexanucleotide repeat expansion in *C9orf72***

Discussion

Our data show that the *C9orf72* hexanucleotide repeat expansion is the most frequent cause of sporadic ALS and sporadic FTD identified thus far, accounting for

***Panel:* Research in context**

Systematic review

We searched Medline up to December, 2011, without language restrictions for relevant publications and selected studies that reported the GGGGCC hexanucleotide repeat expansion in *C9orf72* in pathogenesis of amyotrophic lateral sclerosis (ALS) or frontotemporal dementia (FTD). On the basis of these criteria, seven studies were identified for further assessment (appendix). The number of patients screened for the pathogenic repeat expansion and the phenotype and ethnic origin reported by these studies are summarised in the appendix.

Interpretation

We report the frequency of the *C9orf72* repeat expansion in a large cohort of patients with sporadic ALS and sporadic FTD. We also screened a large number of non-white patients for the expansion, and present frequency data for the mutation in these populations. We confirmed that the *C9orf72* repeat expansion explains a substantial proportion of sporadic ALS (~7·0%) and sporadic FTD (~6·0%) cases in white populations. We also noted that patients with sporadic and familial disease carrying the expansion share a founder risk haplotype, suggesting that these patients have a common ancestor and that the original mutational event that led to the repeat expansion occurred only once in the past. We provide initial estimates of age-related penetrance, showing that 50% of carriers manifest disease by 58 years of age, and that the mutation is fully penetrant by 80 years of age.

about 5·0–7·0% of cases in white Europeans, Americans, and Australians in our large cohort. These frequency rates were slightly higher than were estimates from smaller cohorts obtained at one institution.[9] Before identification of the genetic lesion underlying chromosome 9-linked ALS and FTD, mutations in the *SOD1* gene were the most common known genetic cause of sporadic ALS (accounting for 0·7% of cases in a population-based cohort),[3] whereas mutations in the *PGRN* gene were the most common known cause of sporadic FTD (3·0–4·0% in clinic referral series).[21] The high frequency of the pathogenic expansion in our cohort is consistent with previous genome-wide association studies that identified the association signal on chromosome 9p21 as the only replicable locus in the sporadic form of ALS and FTD.[16,22–24] Our findings confirm the importance of genetics in the pathogenesis of the idiopathic form of these neurodegenerative diseases.

Our haplotype data suggest that the pathogenic GGGGCC hexanucleotide repeat expansion in *C9orf72* arose from a one-off mutation event[16,17] that occurred about 1500 years ago. The geographical distribution of the mutation suggests that the mutation appeared in northern Europe and spread from there. Alternatively, the high frequencies in Finland and other isolated populations could be explained by the history of these communities. Finland and Sardinia are comparatively isolated regions, and have genetically homogeneous populations that originated from a small number of founders.[25] Genetic drift has had a large influence on allele frequencies in these populations and could explain the high occurrence of the mutation in these geographical isolates.

Recognition that all patients carrying the *C9orf72* repeat expansion share a common ancestor has important implications for the interpretation of global frequency data for this mutation. Although the hexanucleotide repeat expansion is common in white Europeans, it is also present in black and Hispanic populations in the USA and individuals from Israel. This finding probably reflects the scale and nature of past human migration and intermarriage between ethnic groups. Similarly, the relative absence of the pathogenic hexanucleotide repeat in India, Asia, and the Pacific Islands might be explained by the greater physical distances of these regions from Europe, and the consequent lack of admixture between populations. Notably, the one Japanese patient who we identified as a carrier of the *C9orf72* expansion carried the Finnish risk haplotype, reinforcing the notion that the expansion occurred on one occasion in the past.

The sharing of a common risk haplotype in the *C9orf72* region of chromosome 9p21 in patients with sporadic and familial ALS suggests that these apparently sporadic cases are actually cryptically related familial cases. This scenario might have occurred for several reasons, including unfamiliarity with the pedigree on the part of the patient or neurologist or because previous generations might have died at a young age before onset of neurological symptoms. The median age at onset in patients with the expansion was 57 years, and life expectancy in the USA began to exceed this point only in the early 1940s.[26] Furthermore, the incomplete penetrance of the mutation, in which not all individuals carrying the expansion manifest a clinical phenotype, might be a contributing factor in apparently sporadic disease. Indeed, we have reported symptom onset in the ninth decade of life in patients carrying the expansion and also encountered two elderly, neurologically healthy individuals with the expansion. Thus, the penetrance of this mutation seems to be complete only at a late stage of life, which is an observation of particular relevance for genetic counselling of healthy individuals carrying the expansion. The molecular biological substrate underlying this variability in age at onset is unclear: it might be driven by differences in expansion lengths between patients, by age-related methylation across the locus, or by genetic factors elsewhere in the genome.

We compared our results with those of previous studies that reported the frequency of the *C9orf72* hexanucleotide repeat expansion in the pathogenesis of ALS and FTD (panel). Data were available from seven studies (appendix). Our study screened one of the largest cohorts of cases of ALS and FTD assessed to date, and also provides an initial report of the frequency of the pathogenic repeat expansion in non-white patients, a detailed examination of the haplotype across the locus, and an initial estimate of age-related disease

penetrance in a large group of individuals carrying the expansion.

Our data have implications for the clinical care of patients diagnosed with ALS and FTD. The clinical standard of care is to offer genetic testing to patients reporting a family history of ALS or FTD,[27] and to reassure patients classified as having sporadic disease that their relatives are not at increased risk of neurodegeneration. On the basis of an analysis of 191 Irish patients with ALS, Byrne and colleagues[28] suggested that genetic testing for the *C9orf72* repeat expansion is unnecessary in affected individuals without a family history of disease or substantial cognitive impairment. By contrast, we believe that genetic testing is a valuable technique for accurate diagnosis of the two disorders and in the decision-making process for patients and their families. The discrepancy between these two views might stem from differences in how sporadic and familial disease were defined in the two studies. Accumulation of sufficient data is an important step towards answering this key question for management of patients. In view of the large number of patients who carry the repeat expansion, investigators and clinicians should at least consider a focused debate on this issue.

Our paper has some limitations. First, the number of patients from some geographical regions was small and the mutational frequencies might change for those ethnic groups as additional patients are screened. Nevertheless, our data for more than 5000 patients with ALS or FTD provide a reasonable estimation of *C9orf72* global frequency. Second, although we have examined the chromosome 9p21 haplotype in a large and diverse cohort of individuals carrying the pathogenic expansion, additional testing of carriers might reveal other haplotypes, thereby indicating that the expansion arose on more than one occasion. Nevertheless, our data suggest that most expansion carriers share a common ancestor.[16,17] Third, we generated age-related penetrance estimates on the basis of data from retrospective cohorts, which potentially leads to overestimation of penetrance. Additional prospective studies examining family kindreds are necessary to confirm these estimates. Finally, case classification as familial or sporadic was done on the basis of clinical questioning at sample collection. The level of scrutiny might have varied between centres and countries, but re-collection of this information for existing cohorts was not feasible.

Contributors

EM, AER, KM, NN, AW, SR, JSS, YA, JOJ, DGH, SA, and JK did laboratory-based experiments and data analysis, and revised the report. ED, MSe, RP, RWO, KCS, HH, JDR, KEM, HP, KT, OA, MSa, GM, MC, FG, ACa, EE, GB, GLF, AMR, HL, LM, VED, and CD collected data from and characterised patients, and revised the manuscript. MAN analysed the data and revised the report. SM, JQT, VMVD, GDS, C-SL, T-HY, HI, YT, ST, ILB, AB, and PS supervised laboratory-based experiments, and revised the report. ACh, GR, JvS, NW, JH, PJT, PH, HRM and SP-B designed the study, supervised laboratory-based experiments, and revised the report. BJT designed the study, supervised laboratory-based experiments, did the data analysis, and drafted the report. The Chromosome 9-ALS/FTD Consortium, The French research network on FTLD/FTLD/ALS, and The ITALSGEN Consortium provided data and helped with data analysis.

Conflicts of interest

PT, PH, HW, SP-B, and BT have a patent pending on the clinical testing and therapeutic intervention for the hexanucleotide repeat expansion of *C9orf72*. JR is Director of the Packard Center for amyotrophic lateral sclerosis Research at Johns Hopkins (MD, USA). All other authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the Intramural Research Programs of the US National Institutes of Health (NIH), National Institute on Aging (Z01-AG000949-02), and National Institute of Neurological Disorders and Stroke (NINDS). The work was also supported by the Packard Center for ALS Research at Hopkins (BJT), the ALS Association (BJT, ACh), Microsoft Research (BJT, PJT), AriSLA (BJT, ACh, MSa), Hersenstichting Nederland Fellowship project B08.03 and the Neuroscience Campus Amsterdam (JS-S), Nuts Ohra Fonds (JvS), Stichting Dioraphte (JvS; grant 09020300), the UK Motor Neurone Disease Association (HM [Motor Neurone Disease Association grant 6057], JH, RWO, KEM, PJS MNDA Grant 6700/3), The Medical Research Council UK (JH, HH, SP-B), the Wellcome Trust (JH, HH, PJS; 069388/z/02/z), The Oxford National Institute for Health Research Biomedical Research Centre (OA), the Helsinki University Central Hospital, the Finnish Academy (PJT), the Finnish Medical Society Duodecim, Kuopio University, the Italian Health Ministry (Ricerca Sanitaria Finalizzata 2007 to ACh), Fondazione Vialli e Mauro ONLUS (ACh), Federazione Italiana Giuoco Calcio (ACh, MSa, BJT) and Compagnia di San Paolo (ACh, GR), the French Agency for Research (ANR-08-MNPS-009-01; AB and ILB), France Alzheimer–Union Nationale des Associations Alzheimer (ILB) and Institut de France Subvention de la Fondation Thierry et Annick DESMAREST (ILB), and the European Community's Health Seventh Framework Programme under grant agreements 259867 (ACh, JK, PJS, MS, CD), Deutsche Forschungsgemeinschaft (MSe; grant SFT.581, TP4). DNA samples for this study were obtained in part from the NINDS repository at the Coriell Cell Repositories (NJ, USA), and from the Australian Motor Neuron Disease DNA Bank, which is funded by National Health and Medical Research Council grant 402703. We thank the DNA extraction and storage facility of the NIH and Welfare/FIMM, Helsinki, Finland and the Institute for Ageing and Health, Campus for Ageing and Vitality, Newcastle University, Newcastle upon Tyne, UK, for their help in extraction of DNA from patients with amyotrophic lateral sclerosis; and also the patients and research participants who contributed samples for this study.

References

1 Logroscino G, Traynor BJ, Hardiman O, et al. Incidence of amyotrophic lateral sclerosis in Europe. *J Neurol Neurosurg Psychiatry* 2010; **81:** 385–90.

2 Valdmanis PN, Daoud H, Dion PA, Rouleau GA. Recent advances in the genetics of amyotrophic lateral sclerosis. *Curr Neurol Neurosci Rep* 2009; **9:** 198–205.

3 Chiò A, Traynor BJ, Lombardo F, et al. Prevalence of SOD1 mutations in the Italian ALS population. *Neurology* 2008; **70:** 533–37.

4 Guerreiro RJ, Schymick JC, Crews C, Singleton A, Hardy J, Traynor BJ. TDP-43 is not a common cause of sporadic amyotrophic lateral sclerosis. *PLoS One* 2008; **3:** e2450.

5 Lai SL, Abramzon Y, Schymick JC, et al. FUS mutations in sporadic amyotrophic lateral sclerosis. *Neurobiol Aging* 2011; **32:** 550.

6 Dion PA, Daoud H, Rouleau GA. Genetics of motor neuron disorders: new insights into pathogenic mechanisms. *Nat Rev Genet* 2009; **10:** 769–82.

7 Ratnavalli E, Brayne C, Dawson K, Hodges JR. The prevalence of frontotemporal dementia. *Neurology* 2002; **58:** 1615–21.

8 Neumann M, Sampathu DM, Kwong LK, et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* 2006; **314:** 130–33.

9 Dejesus-Hernandez M, Mackenzie IR, Boeve BF, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of *c9orf72* causes chromosome 9p-linked FTD and ALS. *Neuron* 2011; **72:** 245–56.

10 Renton AE, Majounie E, Waite A, et al. A hexanucleotide repeat expansion in *C9ORF72* is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 2011; **72:** 257–68.

11 Simón-Sánchez J, Dopper EGP, Cohn-Hokke PE, et al. The clinical and pathological phenotype of *C9orf72* hexanucleotide repeat expansions. *Brain* 2012; published online Feb 2. DOI:10.1093/brain/awr353.

12 Snowden JS, Rollinson S, Thompson JC, et al. Distinct clinical and pathological characteristics of frontotemporal dementia associated with *C9ORF72* mutations. *Brain* 2012; published online Feb 2. DOI:10.1093/brain/awr355.

13 Brooks BR. El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial "Clinical limits of amyotrophic lateral sclerosis" workshop contributors. *J Neurol Sci* 1994; **124** (suppl)**:** 96–107.

14 Clinical and neuropathological criteria for frontotemporal dementia. The Lund and Manchester Groups. *J Neurol Neurosurg Psychiatry* 1994; **57:** 416–18.

15 Traynor BJ, Nalls M, Lai SL, et al. Kinesin-associated protein 3 (KIFAP3) has no effect on survival in a population-based cohort of ALS patients. *Proc Natl Acad Sci USA* 2010; **107:** 12335–38.

16 Laaksovirta H, Peuralinna T, Schymick JC, et al. Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. *Lancet Neurol* 2010; **9:** 978–85.

17 Mok K, Traynor B, Schymick J, et al. The chromosome 9 ALS and FTD locus is probably derived from a single founder. *Neurobiol Aging* 2011; published online Aug 5. DOI:10.1016/j.neurobiolaging.2011.08.005.

18 Reeve JP, Rannala B. DMLE+: Bayesian linkage disequilibrium gene mapping. *Bioinformatics* 2002; **18:** 894–95.

19 Bender BU, Eng C, Olschewski M, et al. VHL c.505 T>C mutation confers a high age-related penetrance bu no increased overall mortality. *J Med Genet* 2001; **38:** 508–14.

20 Passant U, Gustafson L, Brun A. Spectrum of frontal lobe dementia in a Swedish family. *Dementia* 1993; **4:** 160–62.

21 Le Ber I, van der Zee J, Hannequin D, et al. Progranulin null mutations in both sporadic and familial frontotemporal dementia. *Hum Mutat* 2007; **28:** 846–55.

22 Van es MA, Veldink JH, Saris CG, et al. Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat Genet* 2009; **41:** 1083–87.

23 Shatunov A, Mok K, Newhouse S, et al. Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. *Lancet Neurol* 2010; **9:** 986–94.

24 Van Deerlin VM, Sleiman PM, Martinez-Lage M, et al. Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nat Genet* 2010; **42:** 234–39.

25 Kristiansson K, Naukkarinen J, Peltonen L. Isolated populations and complex disease gene identification. *Genome Biol* 2008; **9:** 109.

26 US Census Bureau. Statistical abstract of the United States: 2012, 131st edn. Washington, DC, USA; US Census Bureau, 2011.

27 The EFNS task force on diagnosis and management of amyotrophic lateral sclerosis. EFNS guidelines on the clinical management of amyotrophic lateral sclerosis (MALS)—revised report of an EFNS task force. *Eur J Neurol* 2011; published online Sept 14. DOI:10.1111/j.1468-1331.2011.03501.x.

28 Byrne S, Elamin M, Bede P, et al. Cognitive and clinical characteristics of patients with amyotrophic lateral sclerosis carrying a *C9orf72* repeat expansion: a population-based cohort study. *Lancet Neurol* 2012; **11:** 232–40.

C9ORF72 expansions, parkinsonism, and Parkinson disease

A clinicopathologic study

Johnathan Cooper-Knock, BA*
Antonina Frolov*
J. Robin Highley, DPhil*
Gavin Charlesworth, BSc
Janine Kirby, PhD
Antonio Milano, PhD
Judith Hartley
Paul G. Ince, MD
Christopher J. McDermott, PhD
Tammaryn Lashley, PhD
Tamas Revesz, MD
Pamela J. Shaw, MD
Nicholas W. Wood, PhD‡
Oliver Bandmann, PhD‡

Correspondence to Dr. Bandmann: o.bandmann@sheffield.ac.uk

ABSTRACT

Objective: To determine the histopathologic bases for the observed incidence of parkinsonism in families with *C9ORF72* expansions, which typically cause amyotrophic lateral sclerosis (ALS) and/or frontotemporal dementia.

Methods: DNA was extracted from 377 brains with the histopathologic diagnosis of idiopathic Parkinson disease or related disorders and analyzed for *C9ORF72* expansions. α-Synuclein and p62 immunohistochemistry of the substantia nigra (SN) was undertaken in brains of 17 ALS cases with (*C9ORF72+*) and 51 without (*C9ORF72−*) the *C9ORF72* expansion.

Results: Only 1 of 338 cases with pathologically confirmed idiopathic Parkinson disease had a *C9ORF72* expansion. Similarly, only 1 of 17 *C9ORF72+* brains displayed features suggestive of α-synucleinopathy. In contrast, p62-positive, TDP-43–negative neuronal cytoplasmic inclusions within the SN were considerably more frequent in *C9ORF72+* brain tissue than in the *C9ORF72−* brains ($p = 0.005$). Furthermore, there was a more marked loss of dopaminergic neurons in the SN of *C9ORF72+* ALS brains than *C9ORF72−* ALS brains ($p = 0.029$).

Conclusions: SN involvement is common in *C9ORF72+* ALS but can be clearly distinguished from Parkinson disease–related mechanisms by the presence of p62-positive inclusions and the absence of α-synuclein–positive Lewy bodies or Lewy neurites. ***Neurology***® **2013;81:808–811**

GLOSSARY

ALS = amyotrophic lateral sclerosis; ***C9ORF72*** = chromosome 9 open reading frame 72; **FTD** = frontotemporal dementia; **FTLD** = frontotemporal lobar degeneration; **iPD** = idiopathic Parkinson disease; **PD** = Parkinson disease; **SN** = substantia nigra; **TDP-43** = TAR DNA-binding protein 43.

Substantia nigra (SN) involvement in amyotrophic lateral sclerosis (ALS) has previously been noted clinically[1] and neuropathologically.[2,3] Expansions of *C9ORF72* with >30 repeats (*C9ORF72+*) are the most common identifiable genetic cause of ALS and frontotemporal dementia (FTD).[3,4] We and others have reported parkinsonian phenotypes at a greater frequency within *C9ORF72+* families[3,5] and sporadic cases,[3,6,7] than in those with ALS/FTD who did not have a *C9ORF72* expansion (*C9ORF72−*). Thus, it seems likely that intronic expansions of *C9ORF72* explain at least in part the observed association between ALS and parkinsonism. However, it is currently unclear whether *C9ORF72+* mutation carriers develop parkinsonism due to *C9ORF72+* causing an α-synucleinopathy—as observed in idiopathic Parkinson disease (iPD)—or whether the underlying pathology in these patients is more in keeping with typical *C9ORF72+* extramotor pathology with p62-positive, TDP-43–negative, ubiquitylated neuronal and glial cytoplasmic inclusions.[3] Of note, the 9p21 locus has not been implicated in genetic association studies of iPD.[8]

To further clarify these crucial issues, we have genotyped a large number of brain tissue samples with the histopathologically confirmed diagnosis of iPD or related disorders for *C9ORF72* expansions. We also hypothesized that subclinical involvement of the SN may be more common

Supplemental data at www.neurology.org

*These authors contributed equally to this work.
‡These authors contributed equally to this work.

From the Sheffield Institute for Translational Neuroscience (J.C.-K., J.R.H., J.K., J.H., P.G.I., C.J.M., P.J.S., O.B.), University of Sheffield; Department of Molecular Neuroscience (A.F., G.C., T.L., T.R., N.W.W.), UCL Institute of Neurology, Queen Square, London; and Sheffield Diagnostic Genetic Service (A.M.), Sheffield Children's NHS Foundation Trust, Western Bank, UK.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

in *C9ORF72*+ than in *C9ORF72*− brains of patients who clinically presented only with features in keeping with ALS/FTD.

METHODS **Standard protocol approvals, registrations, and patient consents.** Ethical approval was obtained from the respective local and national ethics committees. Postmortem histology reports were provided by the Parkinson's UK Brain Bank and the Queen Square Brain Bank for Neurological Disorders.

Subjects. A total of 377 cases with a clinical diagnosis of PD were included; Lewy-body–positive, α-synucleinopathy was pathologically confirmed in 338 (90%) of cases (see table). Thus, a sufficient number of cases was obtained to allow comparison with the known frequency of expanded *C9ORF72* in controls.[3]

Genotyping and immunohistochemistry. DNA was extracted from brain tissue using standard methods and analyzed for *C9ORF72* expansions as previously described.[3] α-Synuclein and p62 immunohistochemistry was performed on 17 *C9ORF72*+ cases of ALS, including one case known to have autopsy-confirmed iPD as well as ALS.[3] Immunohistochemistry for p62 was also performed on an additional 51 *C9ORF72*− ALS cases. The SN was examined on one side of each brain. Seven-micron-thick tissue sections from selected blocks were subjected to immunohistochemistry using antibodies to α-synuclein (Novocastra, Milton Keynes, UK) and p62 (BD Transduction Laboratories, Oxford, UK). α-Synuclein pathology was assessed as present or absent in *C9ORF72*+ cases. The number of p62-positive inclusions was classified as high (≥10 positive neuronal cytoplasmic inclusions), intermediate (5–9 inclusions), or low (≤4 inclusions). The number of cases with and without the *C9ORF72* mutation with high, intermediate, and low numbers of p62-positive neuronal cytoplasmic was compared by χ^2. To determine whether the ubiquitylated neuronal cytoplasmic inclusion pathology of the SN was associated with neuronal loss, both *C9ORF72*+ and *C9ORF72*− brains were semiquantitatively assessed as having no, mild, or severe neuronal loss. The extent of the neuronal cell loss was compared by χ^2.

RESULTS Only one of the brain bank iPD cases had a *C9ORF72* expansion containing >30 repeats (1/377 = 0.2% of the total number of cases screened, and 1/338 = 0.3% of the Lewy-body–positive cases). This *C9ORF72*+ patient presented with clinically typical PD at the age of 67 years. His father had died of ALS. Neuropathologic assessment revealed features of 1) classic PD with Braak stage 6, diffuse neocortical Lewy-body pathology; 2) classic TDP-43 pathology with frontotemporal lobar degeneration (FTLD)-TDP type-A features[9]; and 3) *C9ORF72*-ALS/FTLD pathology with numerous p62-positive, TDP-43–negative neuronal cytoplasmic inclusions of star-shaped morphology in the hippocampus, and smaller cytoplasmic inclusions in cerebellar granule cells. Unfortunately, his spinal cord was not available.

All but one of the 17 *C9ORF72*+ ALS brains were devoid of α-synuclein–positive neuronal cytoplasmic inclusions in the SN. The single case with α-synuclein pathology was known to have coincident PD-ALS and has been discussed elsewhere.[3]

The 17 *C9ORF72*+ ALS brains had a considerably higher number of p62-positive, TDP-43–negative neuronal cytoplasmic inclusions in the SN (7 cases with >10, 2 cases with 5–9, and 8 cases with ≤4 p62-positive inclusions) than the 51 *C9ORF72*− ALS cases (4 cases >10, 6 cases with 5–9, and 41 cases with ≤4 p62-positive inclusions; $\chi^2 = 10.724$, $df = 2$, $p = 0.005$). No/moderate/severe neuronal cell loss was observed in 3/9/5 cases with *C9ORF72* mutations and 22/17/4 cases without this mutation ($\chi^2 = 7.074$, $df = 2$, $p = 0.029$). Thus, the burden of p62-positive disease is much greater than that seen on α-synuclein immunohistochemistry and was associated with a variable degree of neuronal loss in the SN (figure).

DISCUSSION Previous studies investigating a possible association among *C9ORF72* expansions, parkinsonism, and iPD have concentrated on patients with the clinical rather than the pathologic diagnosis of iPD.[e1–e7] These studies have all concluded that *C9ORF72* expansions are not a common cause of iPD; however, given that clinical diagnosis has a higher false-positive rate, we chose to conduct a study of pathologically confirmed iPD. Furthermore, by focusing on neuropathology, we also investigated the pathologic basis of parkinsonian presentations in *C9ORF72*+ ALS patients.

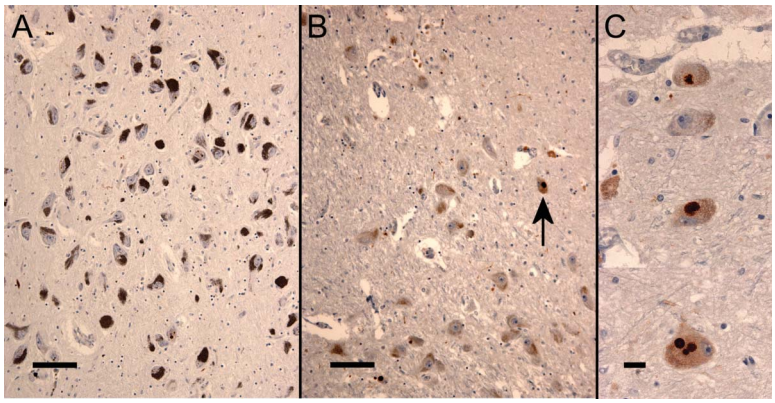
We identified a single *C9ORF72*+ patient with clinically typical iPD of 377 tested; this frequency is similar to that in controls[3] and thus we conclude that *C9ORF72* expansions are not a major cause of iPD. Notably, this patient had a family history of ALS and neuropathology consistent with *C9ORF72*+ disease; although it was not possible to investigate motor neuron loss in the spinal cord, we suspect this patient had subclinical FTD. The presence of type-A FTLD

Table Characteristics of tissue samples

Biobank	No. of cases	Male: female	Mean age (range) at onset, y	Family history of parkinsonism, n (%)	Cognitive impairment, n (%)	Lewy bodies present, n (%)
UK Parkinson's Disease Society Tissue Brain Bank	141	2.4:1	65.2 (35–86)	21 (25)	98 (70)	114 (81)
Queen Square Brain Bank	236	1.5:1	63.8 (30–85)	NA	120 (51)[a]	224 (95)

Abbreviation: NA = not available.
[a] Information about cognitive impairment was not available for postmortem tissue of 62 patients in the Queen Square Brain Bank for Neurological Disorders.

Figure Substantia nigra histopathology in *C9ORF72*− and *C9ORF72*+ cases



Photomicrographs of the substantia nigra after immunohistochemistry for p62 in *C9ORF72*− brains (A) and *C9ORF72*+ brains (B, low power, and C, high power) of patients with the clinical diagnosis of amyotrophic lateral sclerosis showing neuronal loss and p62-positive cytoplasmic inclusions (arrow) in the *C9ORF72*+ brain. Bar = 100 μm (A and B), and 20 μm (C).

pathology suggests that motor neuron pathology was unlikely.

It is noteworthy that of the 2 cases that were known to have coincident PD-ALS in the Sheffield brain bank, one did[3] and one did not have the *C9ORF72* repeat expansion.

The absence of α-synuclein pathology in the SN of the vast majority of *C9ORF72*+ brains further strengthens our assumption that the intracellular mechanisms leading to neuronal cell loss in ALS/FTD and those causing α-synuclein pathology in iPD are distinct. In contrast, p62-positive, TDP-43–negative inclusions in combination with neuronal loss are considerably more common in the SN of *C9ORF72*+ ALS patients than in *C9ORF72*− ALS. This p62-positive extrapyramidal pathology is therefore the likely cause of the previously reported increased incidence of parkinsonian features in *C9ORF72*-related ALS. One could therefore consider *C9ORF72*-related neurodegeneration as a clinically and pathologically heterogeneous syndrome characterized by a combination of TDP-43 proteinopathy with superimposed extramotor p62-positive, TDP-43–negative pathology. The distribution and severity of this latter pathology is likely to govern the presence of cognitive impairment (in the presence of hippocampal and neocortical pathology) or parkinsonism (in the presence of basal ganglia pathology).

Until the pathogenesis of *C9ORF72* disease is fully understood, it remains impossible to exclude *C9ORF72* expansions as a very rare cause of α-synucleinopathy and clinical iPD. However, our observation of an alternative pathologic basis for the observed incidence of parkinsonism in *C9ORF72*+ patients, significantly strengthens the case that *C9ORF72* disease and α-synucleinopathy represent distinct pathologic entities.

Understanding that *C9ORF72* expansions are a cause of both ALS and a parkinsonian phenocopy is likely to be crucial to the counseling and management of patients with ALS presenting with parkinsonian features, particularly if they have a family history of ALS/PD. Genetic testing for expansions of *C9ORF72* will help to differentiate patients with *C9ORF72* neurodegeneration from those who have developed more typical PD; a similar suggestion has been made for the use of *C9ORF72* genotyping in the differentiation of true Alzheimer disease from FTD caused by mutation of *C9ORF72*.[10]

AUTHOR CONTRIBUTIONS

The study was conceived and designed by authors J.C.-K., J.R.H., G.C., P.J.S., N.W.W., and O.B. Data acquisition was performed by authors J.C.-K., A.F., J.R.H., G.C., J.K., A.M., J.H., S.B.W., P.G.I., C.M.D., T.L., and T.R. Data analysis and interpretation were performed by J.C.-K., A.F., J.R.H., G.C., T.L., T.R., N.W.W., and O.B. The manuscript was critically revised by J.C.-K., J.R.H., P.J.S., and O.B. The study was supervised by T.R., P.J.S., N.W.W., and O.B.

ACKNOWLEDGMENT

The authors are grateful to all of the patients with ALS and PD who donated biosamples for research purposes.

STUDY FUNDING

The biosample collection for the ALS cases was supported by the MND Association and the Wellcome Trust (P.J.S.). Genetic testing was supported by National Institute for Health Research (NIHR) CLAHRC for South Yorkshire and by a Wellcome Trust/MRC joint strategic award (WT089698/Z/09/Z) (N.W.W.). The Queen Square Brain Bank for Neurological Disorders is supported by the Reta Lila Weston Institute for Neurological Studies, the MRC, and the PSP (Europe) Association. The research was, in part, supported by the NIHR Biomedical Research Unit in Dementia based at University College London Hospitals, University College London (UCL).

DISCLOSURE

J. Cooper-Knock is supported by MND Association/Medical Research Council Lady Edith Wolfson Fellowship awards (G0 800380) and (R/132205). A. Frolov reports no disclosures. J. Highley is supported by MND Association/Medical Research Council Lady Edith Wolfson Fellowship awards (G0 800380) and (R/132205), and by Joint Programme for Neurodegenerative Disease (JPND) grant SOPHIA. G. Charlesworth reports no disclosures. J. Kirby is supported by JPND grant SOPHIA and an FP7 grant EuroMOTOR (no. 259867). A. Milano, J. Hartley, P. Ince, and C. McDermott report no disclosures. T. Lashley is supported by an Alzheimer's Research UK fellowship award (ARUK-RF2012-1). T. Revesz received research support from Alzheimer's Research UK, Parkinson's UK, and the Multiple System Atrophy Trust, received honoraria for lectures from Novartis, serves as editorial board member of *Acta Neuropathologica*, as editorial advisory board member of *Neuropathology* and *Applied Neurobiology*, and as Associate Editor of the *Journal of Parkinson's Disease*. He also received funding for 2 trips from Merck Serono. P. Shaw is supported by JPND grant SOPHIA and an FP7 grant EuroMOTOR (no. 259867). N. Wood is supported by a Wellcome Trust/MRC joint strategic award (WT089698/Z/09/Z). O. Bandmann is supported by Parkinson's UK (G-1007). He received a speaker honorarium from GlaxoSmithKline in 2012. He is a member of the Editorial Board of *Neurology*®.

Received February 15, 2013. Accepted in final form May 24, 2013.

REFERENCES

1. Lim YM, Park HK, Kim JS, et al. Clinical and neuroimaging characteristics in neurodegenerative overlap syndrome. Neurol Sci 2013;34:875–881.

2. Nishihira Y, Tan CF, Onodera O, et al. Sporadic amyotrophic lateral sclerosis: two pathological patterns shown by analysis of distribution of TDP-43-immunoreactive neuronal and glial cytoplasmic inclusions. Acta Neuropathol 2008;116:169–182.
3. Cooper-Knock J, Hewitt C, Highley JR, et al. Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. Brain 2012;135:751–764.
4. Majounie E, Renton AE, Mok K. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. Lancet Neurol 2012;11:323–330.
5. Renton AE, Majounie E, Waite A, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 2011;72:257–268.
6. Boeve BF, Boylan KB, Graff-Radford NR, et al. Characterization of frontotemporal dementia and/or amyotrophic lateral sclerosis associated with the GGGGCC repeat expansion in C9ORF72. Brain 2012;135:765–783.
7. O'Dowd S, Curtin D, Waite AJ, et al. C9ORF72 expansion in amyotrophic lateral sclerosis/frontotemporal dementia also causes parkinsonism. Mov Disord 2012;27:1072–1074.
8. Lill CM, Roehr JT, McQueen MB, et al. Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: the PDGene database. PLoS Genet 2012;8:e1002548.
9. Mackenzie IR, Neumann M, Baborie A, et al. A harmonized classification system for FTLD-TDP pathology. Acta Neuropathol 2011;122:111–113.
10. Majounie E, Abramzon Y, Renton AE, et al. Repeat expansion in C9ORF72 in Alzheimer's disease. N Engl J Med 2012;366:283–284.

RESEARCH PAPER

Concurrence of multiple sclerosis and amyotrophic lateral sclerosis in patients with hexanucleotide repeat expansions of *C9ORF72*

Azza Ismail,[1,3] Johnathan Cooper-Knock,[1,3] J Robin Highley,[1] Antonio Milano,[7] Janine Kirby,[1] Emily Goodall,[1] James Lowe,[5] Ian Scott,[5] Cris S Constantinescu,[4] Stephen J Walters,[2] Sian Price,[3] Christopher J McDermott,[1,3] Stephen Sawcer,[6] D Alastair S Compston,[6] Basil Sharrack,[1,3] Pamela J Shaw[1,3]

[1]Academic Neurology Unit and Department of Neuroscience, Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, Sheffield, UK
[2]School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK
[3]Department of Neurology, Academic Directorate of Neuroscience, Royal Hallamshire Hospital, Sheffield, UK
[4]Department of Neurology, University of Nottingham and Queen's Medical Centre, Nottingham, UK
[5]Department of Neuropathology, University of Nottingham and Queen's Medical Centre, Nottingham, UK
[6]Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK
[7]Sheffield Diagnostic Genetic Service, Sheffield Children's NHS Foundation Trust, Western Bank, Sheffield,UK

Correspondence to
Professor Pamela J Shaw, Academic Neurology Unit, Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, 385A Glossop Road, Sheffield S10 2HQ, UK; pamela.shaw@sheffield.ac.uk

Received 28 May 2012
Revised 11 September 2012
Accepted 11 September 2012
Published Online First 20 October 2012

ABSTRACT

Background Crossover in the pathogenic mechanisms of amyotrophic lateral sclerosis (ALS) and multiple sclerosis (MS) has been described but is poorly understood. A GGGGCC hexanucleotide repeat expansion of *C9ORF72* has recently been identified in a significant proportion of patients with ALS.

Methods In approximately 650 patients diagnosed with ALS from the North of England we identified seven patients who initially presented with MS. DNA obtained from five patients with MS-ALS and 215 patients with MS alone was screened for the *C9ORF72* expansion. Post-mortem material was examined from two patients with MS-ALS. Gene expression profiling was performed on lymphoblastoid cells and levels of CXCL10 were measured in cerebrospinal fluid (CSF) from patients with ALS with and without the *C9ORF72* expansion and controls.

Results Concurrence of MS and ALS is higher than expected in our population. The *C9ORF72* expansion was identified in 80% of patients with MS-ALS but not in those with MS alone. In the presence of preceding MS, *C9ORF72*-ALS was more rapidly progressive. MetaCore analysis identified alteration of the NF-κB pathway in *C9ORF72*-ALS and non-*C9ORF72*-ALS. NF-κB activation is associated with increased expression of the neuroprotective cytokine *CXCL10* but, in *C9ORF72*-ALS, *CXCL10* is downregulated and CSF levels are reduced.

Conclusions We propose that MS-associated neuroinflammation may affect penetrance and progression of the *C9ORF72* expansion. In particular, the NF-κB pathway is activated in MS and appears to be dysfunctional in *C9ORF72*-ALS. Aberrant downregulation of *CXCL10* may explain the predisposition of *C9ORF72* expansion carriers to develop ALS in the context of MS and NF-κB activation, and offers a potential therapeutic target.

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is characterised by rapidly progressive degeneration of the motor neurons with consequent progressive failure of the neuromuscular system and eventual death from respiratory compromise, typically within 3 years from symptom onset. Clear Mendelian inheritance is present in approximately 5% of ALS cases.[1] Multiple sclerosis (MS) is an immune mediated disorder of the CNS in which inflammation results in secondary degenerative pathology leading to progressive disability.[2] Although these diseases have seemingly distinct characteristics, there are several reports describing concurrent ALS and MS.[3–8] Until now this association has been largely unexplained.

It has recently been demonstrated that a large proportion of ALS is related to a GGGGCC hexanucleotide repeat expansion in intron 1 of *C9ORF72*.[9–11] In our North of England population, the expansion occurs in 7% of apparently sporadic ALS cases and 43% of familial ALS.[11] Significant variability in the phenotype of *C9ORF72* ALS has been reported,[11] and therefore a role for *C9ORF72* expansions in patients with MS-ALS was explored. DNA from five patients with MS-ALS and 215 patients with MS alone was examined for expansions of *C9ORF72*. Where available, post-mortem material was used to confirm the clinical diagnoses.

Expansions of *C9ORF72* were strongly associated with MS-ALS but absent in patients with MS alone. In addition, concurrent MS was associated with a more rapidly progressive phenotype than pure *C9ORF72*-ALS. Understanding the interaction between MS and *C9ORF72*-ALS has the potential to identify novel therapeutic targets for the treatment of *C9ORF72*-ALS in MS and more broadly. Pathways related to NF-κB have been linked to ALS[12] and to neurodegeneration in MS,[13] and appear to be important in *C9ORF72*-ALS.

METHODS

Identification of cases with concurrent MS and ALS

Patients with MS who subsequently developed ALS were prospectively identified between 2006 and 2012. The clinical case notes of the identified patients were accessed and reviewed. All patients were under the care of senior consultant neurologists and diagnosed with MS and ALS as defined by the modified McDonald's[14] and El Escorial criteria,[15] respectively.

Extraction of DNA from patients with concurrent MS and ALS

The study was approved by the South Sheffield Research Ethics Committee and informed consent was obtained for all samples. DNA from five patients with MS-ALS was obtained from the

Sheffield MND Blood DNA Biobank. DNA was extracted from blood using the Nucleon Blood and Cell Culture Genomic Extraction kit (Tepnel, UK) according to the manufacturer's protocol. All patients were UK Caucasians.

Extraction of DNA from patients with MS

The study was approved by the Thames Valley Multi-Centre Research Ethics Committee and informed consent was obtained for all samples. DNA from 215 patients with MS alone was obtained from the Cambridge MS DNA Bank. These patients had standard demographic characteristics with a female:male ratio of 2.96 : 1, average age of onset 31.8 years (range 14–60 years) and average disease duration 16.4 years (range 1–57 years); 15.9% had a family history of MS, 14.5% had primary progressive disease and 85.5% had relapsing remitting or secondary progressive disease.

Screening for the *C9ORF72* hexanucleotide repeat sequence by repeat primed PCR

Genomic DNA (100 ng) was amplified using the primers and method described previously.[11] Fragments were analysed on an ABI3730 capillary analyser (Applied Biosystems, Life Technologies Corporation, California, USA) using a 60 s injection time. Fragment data were analysed using Peak Scanner Software (Applied Biosystems, Life Technologies Corporation).

Neuropathological evaluation

Where available, neuropathology was evaluated as described previously.[11]

Samples for gene expression profiling

Lymphoblastoid cell lines were obtained from patients with familial ALS from the Wellcome Trust/Motor Neurone Disease Association-funded ALS/MND DNA bank and associated lymphoblastoid cell line repository in the UK. Samples included 10 patients with ALS with the hexanucelotide expansion in *C9ORF72*, 16 patients without the expansion and 10 non-related age- and sex-matched spouse controls. No samples were available from patients with MS-ALS.

RNA preparation and microarray analysis

RNA was extracted from lymphoblastoid cells as previously described.[16] The quality (2100 bioanalyzer, RNA 6000 Pico LabChip; Agilent, California, USA) and quantity (NanoDrop 1000 spectrophotometer) of the RNA from all of the samples were assessed. All RNA used in gene expression profiling had a RIN >7.8. 150 ng from each sample was linearly amplified using One Cycle Amplification protocol (Affymetrix). Fifteen micrograms of amplified cRNA from each case was fragmented and each hybridised individually to Human Genome U133 Plus 2.0 GeneChips (Affymetrix), according to the manufacturers' protocols. Following stringency washes, chips were stained and scanned and GeneChip Operating Software used to produce signal intensities for each transcript.

Statistical analysis of microarray results

Probe set summarisation was carried out using the puma tool.[17] A disease versus control comparison (puma, IPPLR[17 18]) was performed to identify genes differentially expressed between patients with the expansion in *C9ORF72* and controls and other non-*C9ORF72*-ALS. MetaCore (GeneGo Inc) was used to identify likely transcription factor activity based on the differentially expressed genes.

Quantitative polymerase chain reaction for CXCL10 mRNA in lymphoblastoid cells

Total RNA from lymphoblastoid cells was amplified using High Capacity RNA-to-cDNA kit (Applied Biosystems). Quantitative PCR (QPCR) primers for CXCL10 transcripts were designed using Eurofins online primer design software (http://www.eurofinsdna.com). QPCR of the C9ORF72-ALS cases, non-C9ORF72-ALS cases and controls was performed using Brilliant II SYBR Green QPCR Master Mix (Stratagene) on the Stratagene 3000, as described previously.[19] Mann–Whitney U tests were used to determine if the relative differences in *CXCL10* expression in lymphoblastoid cells between C9ORF72-ALS samples and controls, and between C9ORF72-ALS and non-C9ORF72 ALS samples were statistically significant.

Collection of CSF and multiplexed fluorescent bead-based immunoassay of CXCL10

The study was approved by the South Sheffield Research Ethics Committee and informed consent was obtained for all samples. Cerebrospinal fluid (CSF) samples were obtained by lumbar puncture and immediately centrifuged at 800 rpm at 4°C for 5 min. The liquid phase of CSF that excluded the sedimented cells was stored at −80°C until cytokine assay. Levels of CXCL10 were measured using the BD Cytometric Bead Array (CBA). Levels of CXCL10 were detected according to the manufacturer's instructions for BDCBA Flex sets using BDHuman Master Buffer kit. Cytokine concentration was calculated by reference to a standard curve derived from the manufacturer's cytokine standard (0–2500 pg/ml), with the lower limit of detection determined to be 10 pg/ml. Differences between groups were determined by a Mann–Whitney U test.

RESULTS

Identification of patients with concurrent MS and ALS

Between 2006 and 2012 we prospectively identified seven patients with MS who subsequently developed ALS. Five patients presented to the Neurology Department in Sheffield which cares for the majority of patients with MS and almost all patients with ALS in a catchment population of 2.2 million. Two further patients presented to the Department of Neurology in Nottingham and were subsequently referred to the Sheffield MND Clinic. The Nottingham Neurology service also serves a catchment area of approximately 2.2 million. A total of 348 patients were diagnosed with ALS in Sheffield and approximately 300 in Nottingham during the same period. The prevalence of concurrent MS in this series of patients with ALS was therefore 1.1% (95% CI 0.5% to 2.2%) compared with a prevalence of MS in the general population of 0.1%.[20] It is possible that additional cases of MS, ALS and MS-ALS were not identified and therefore perhaps our figure represents a minimum estimate of the strength of association between the pathologies. All patients were serially assessed by neurological examination and underwent appropriate neurophysiological and laboratory investigation and neuroimaging. Patients 1 and 4 had post-mortem examinations (figures 1 and 2 respectively). The patients' demographic and clinical features are summarised in table 1, and their neuroimaging, electromyography (EMG) and CSF findings are summarised in table 2.

Age at onset of ALS in patients with MS-ALS was not significantly different from the whole cohort of patients with ALS, and age at onset in the subset of patients with MS-ALS with the *C9ORF72* expansion was not significantly different from those with pure *C9ORF72*-ALS. Mean age at onset in the

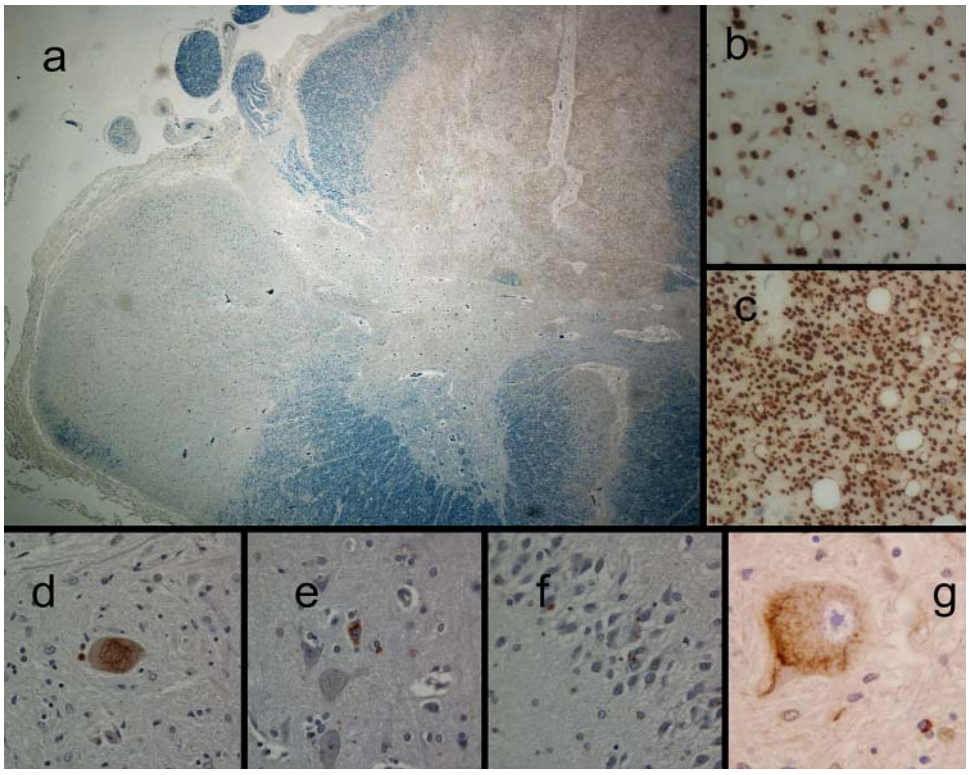


Figure 1 Neuropathological features of patient 1. A solochrome cyanine (myelin) stain shows demyelination in the lateral descending tracts secondary to upper motor neuron degeneration and a chronic multiple sclerosis plaque in the dorsal columns (A). Immunohistochemistry for neurofilament shows axonal depletion in the lateral descending tract (B) but preservation in the region of demyelination in the dorsal column (C). Immunohistochemistry for p62 reveals ubiquitylated neuronal cytoplasmic inclusions in the anterior horn (D), motor cortex (E) and dentate fascia of the hippocampus (F). Immunohistochemistry for TDP-43 labels neuronal and glial cytoplasmic inclusions in the anterior horn.

patients with MS-ALS was 52.9 years (SD 10.0, range 52–67); specifically, in patients with MS-ALS with the *C9ORF72* expansion the mean age at onset was 59.3 years (SD 6.60, range 52–67). This compares with a mean age at onset of 57.3 years (SD 8.9, range 27–74) in patients with pure *C9ORF72*-ALS and 60.2 years in the overall cohort. Mean disease duration in the patients with MS-ALS was 31.1 months (SD 19.5, range 6–60); specifically, in patients with MS-ALS with the *C9ORF72* expansion, mean disease duration was 24 months. This compares with a mean disease duration of 30.5 months (SD 13.3, range 7–60) in pure *C9ORF72*-ALS and 34.7 months in the overall cohort. There was no statistically significant difference in disease duration between patients with MS-ALS and the overall cohort, but disease duration in patients with MS-ALS with the C9ORF72 expansion was significantly shorter than in the overall cohort ($p=2.26E-14$, $df=393$, $t=7.93$) and in those with pure C9ORF72-ALS ($p=0.0004$, $df=60$, $t=3.74$). In no patient with MS-ALS was there any evidence of cognitive impairment although patient 1 had a family history of frontotemporal dementia. In all patients ALS was preceded by MS. The average age at onset of MS in the *C9ORF72* MS-ALS patients was 37.8 years (range 22–46), which is relatively old but consistent with the fact that two of these patients had primary progressive disease which is associated with a later mean age of onset.[21]

Case histories

Patient 1 developed her first neurological symptoms at the age of 23. Aged 33 years she was diagnosed with relapsing-remitting MS which developed into a secondary progressive phase a few years later. At the age of 62 she developed progressive dysarthria, dysphagia and rapid deterioration of her mobility; at this point she was diagnosed with ALS. Consistent with MS, MRI of the brain showed periventricular white matter lesions and CSF analysis was positive for oligoclonal bands. Consistent with ALS, EMG showed evidence of denervation involving cranial, cervical, thoracic and lumbosacral segments. She died of pneumonia 2 years after onset of the ALS symptoms. At autopsy (figure 1) the brain showed evidence of chronic plaques in a number of regions: (1) deep periventricular frontal and parietal white matter (remyelinated 'shadow' plaques); (2) right cerebral peduncle; (3) left anterior medial basal pons (not including pyramidal fibres); (4) right pyramid of the medulla; and (5) the dorsal columns of the cervical, thoracic and lumbar spinal cord.

Immunohistochemistry for p62 revealed ubiquitylated cytoplasmic inclusions in the glossopharyngeal, dorsal vagal nuclei, hippocampus (including both dentate granule cells and pyramidal cells of CA3 and CA4) and motor cortex. In the spinal cord there was severe loss of motor neurons with Bunina bodies and skein-like ubiquitylated cytoplasmic inclusions in residual neurons. There was severe pallor of myelin staining in the descending lateral and ventral motor tracts with a marked microglial reaction. The patient's daughter was subsequently diagnosed with MS at the age of 41 years and to date shows no features of ALS. The patient's mother had a history of frontotemporal dementia and ALS.

Patient 2 presented with relapsing-remitting MS at the age of 49 years and developed bulbar-onset ALS 3 years later. Widespread evidence of denervation was apparent on EMG. She died 4 years after the diagnosis of ALS. There was no family history of ALS or MS.

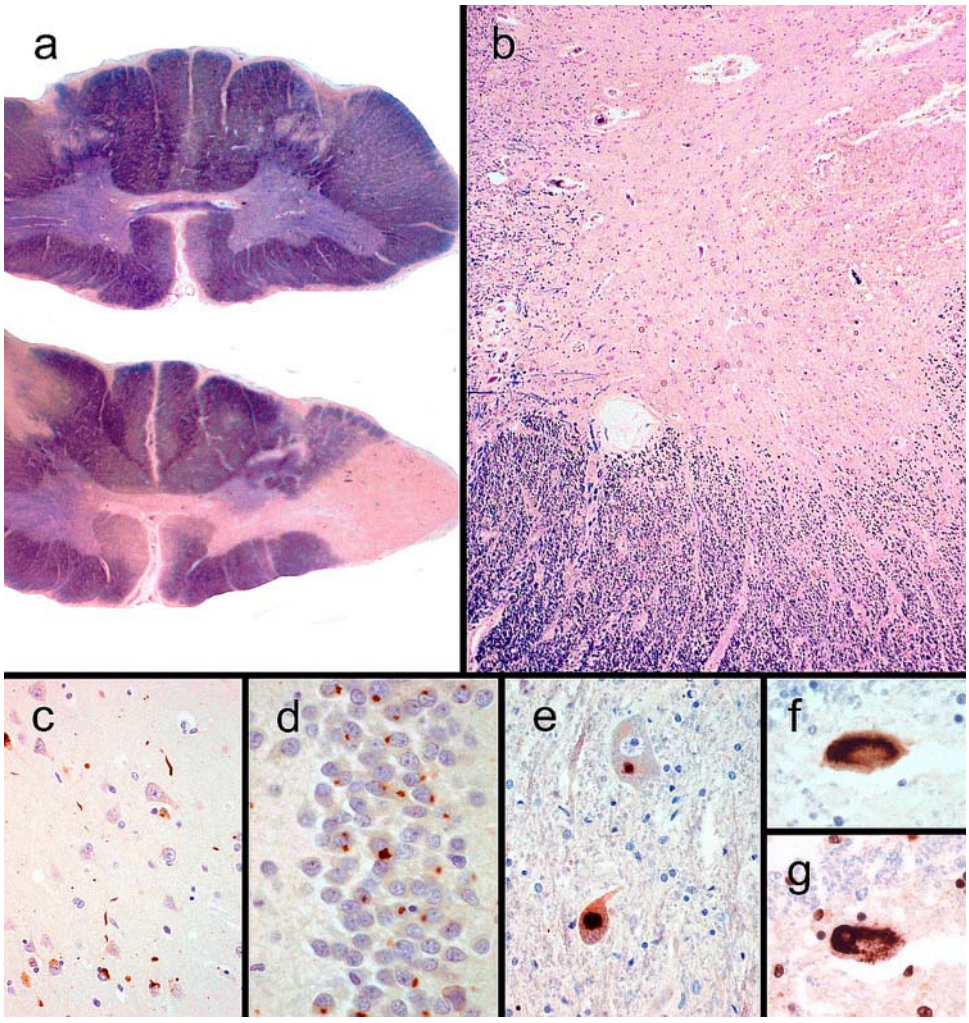


Figure 2 Neuropathological features of patient 4. Myelin-stained (solochrome cyanin) sections of spinal cord show demyelination in the corticospinal tract; the lower image shows additional punched-out plaques of demyelination which includes grey matter demyelination in one anterior horn region (A) and a higher magnification image shows grey matter myelin loss (B). Immunohistochemistry for p62 reveals ubiquitylated neuronal cytoplasmic inclusions in the frontal cortex (C), dentate fascia of the hippocampus (D) and motor neurons of the spinal cord (E, F). Immunohistochemistry for TDP-43 labels neuronal cytoplasmic inclusions in the anterior horn of spinal cord (G).

Patient 3 presented with relapsing-remitting MS at the age of 43 years and was treated with interferon beta 1b. She remained relapse-free for a decade. At the age of 52 she developed lower limb onset ALS and died 2 years later. There was no family history of ALS or MS.

Patient 4 was diagnosed with MS at the age of 46. His disease ran a primary progressive course until he developed ALS after 21 years. He died 2 years later. He had a family history of thyroid disease and persistently elevated thyroid peroxidase antibodies, but no history of other neurological disease. There was no family history of ALS or MS. At autopsy the CNS showed a single 2 mm plaque in the angle of the right lateral ventricle and multiple focal plaques in the spinal cord. Immunohistochemistry revealed ubiquitylated neuronal cytoplasmic inclusions in the dentate fascia of the hippocampus, motor cortex and lower motor neurons.

Patient 5 presented at the age of 39 years with atypical rapid primary progressive MS which failed to respond to mitoxantrone treatment. One year later he developed lower motor neuron signs and was diagnosed with ALS. He died 6 months later. There was no family history of ALS or MS.

Patient 6 presented at the age of 40 years with pain and weakness of her left leg and was diagnosed with MS. Five months later she developed progressive weakness of her arms and legs and was diagnosed with ALS. She suffered a steady decline until she died 5 years later. Her maternal uncle and his daughter both suffered from MS.

Patient 7 developed progressive upper and lower limb weakness on the left side beginning at the age of 56, followed by dysphagia and dysarthria 5 months later. She had a history of hypothyroidism and hypertension. Examination and investigation by EMG confirmed the diagnosis of limb-onset ALS. MRI revealed multiple periventricular lesions extending to the lateral aspect of the corpus callosum consistent with demyelination. Her CSF showed a raised CSF/serum IgG ratio but no oligoclonal bands. There was no family history of neurological disease.

Results of screening for *C9ORF72* hexanucleotide expansion in MS-ALS cases

An expanded number of GGGGCC hexanucleotide repeats within *C9ORF72* was present in all but one (patient 6) of the five patients with MS-ALS (1, 3, 4, 6 and 7) for whom DNA was available for assessment. Our previous work has shown that the expected proportion of *C9ORF72* expansions in patients with pure ALS is 7% in sporadic ALS and 43% in familial ALS.[11] Based on these proportions, the occurrence of a

Table 1 Demographic and clinical features of individual patients with multiple sclerosis-amyotrophic lateral sclerosis (MS-ALS)

Patients	Sex	Age of MS onset (years)	Symptoms at MS onset	Classification of MS	Initial neurological examination	Age of ALS onset (years)	Symptoms of ALS onset	Last neurological examination	Definition of ALS by El Escorial criteria	ALS duration (months)
Patient 1	F	22	Paraesthesia in lower limbs	SP	Pale optic discs, jerky eye movements, absent jaw jerk, incoordination of upper limbs, spastic paraparesis	62	Bulbar	Spastic dysarthria, wasted and fasciculating tongue, brisk jaw jerk, spastic quadraparesis, generalised wasting	Definite ALS	24
Patient 2	F	49	Bilateral lower limb paraesthesia, bladder dysfunction	RR	↓ sensation in right hand, pallor of right optic disc	52	Bulbar	Wasted fasciculating and weak tongue, weak neck flexion and extension, spastic quadraparesis, ↓ vibration sense in lower limbs	Definite ALS	56
Patient 3	F	43	Diplopia, parasthesia in lower limbs	RR	Left disc pallor, nystagmus, weakness of right upper limb, brisk tendon reflexes, ↓ pinprick sensation in right hand	52	Lower limb	Right temporal pallor, nystagmus, ↓ palatal movement, spastic wasted tongue, brisk jaw reflex, global wasting and weakness in upper limbs, spastic tetraparesis, proximal fasciculations, tendon reflexes brisk, extensor planters, ↓ pinprick sensation on left foot	Definite ALS	24
Patient 4	M	46	Weakness of right lower limb	PP	Spastic paraparesis, bilateral ↓vibration sense lower limbs	67	Bulbar	Wasted and fasciculating tongue, spastic quadraparesis, ↓ vibration sense in lower limbs	Definite ALS	24
Patient 5	M	39	Difficulty walking, parasthesia in left hand	PP	Left disc pallor, bilateral nystagmus, internuclear ophthalmoplegia, bilateral limb ataxia, spastic quadraparesis	40	Bulbar	Weak spastic tongue, spastic tetraparesis, widespread fasciculations	Definite ALS	6
Patient 6	F	40	Pain and weakness of left lower limb	Atypical PP	Wasted fasciculating tongue, fasciculations, wasting of IHM, spastic quadraparesis, no sensory deficit	41	Lower limb	Wasted fasciculating tongue, fasciculations, wasting of IHM, spastic quadraparesis, no sensory deficit	Definite ALS	60
Patient 7	F	–	–	–	–	56	Lower limb	Spastic dysarthria and dysphonia, weakness neck extension and flexion, wasted, fasciculating and spastic tongue, brisk jaw reflex, wasting of IHM, widespread fasciculations, spastic tetraparesis, normal sensation	Definite ALS	24

ALS,amyotrophic lateral sclerosis; IHM, intrinsic hand muscles; MS, multiple sclerosis; PP, primary progressive; RR, relapsing remitting; SP, secondary progressive.

C9ORF72 expansion in four out of five of our MS-ALS cases (one familial and four sporadic ALS) by chance is extremely unlikely (OR 3.27, p<0.001). This strongly suggests that the co-occurrence of MS in a patient with ALS is driven in some way by expansion of *C9ORF72*.

Results of screening for the *C9ORF72* expansion in MS cases

No *C9ORF72* expansions, as defined by >30 hexanucleotide repeats,[9 10] were found in the DNA of any patient from a cohort of 215 patients with MS. The median number of repeats in the pure MS cases was 4 (range 1–22). This is not significantly different from the number of repeats found in controls.[11]

Results of gene expression profiling of lymphoblastoid cells from *C9ORF72*-ALS, non-*C9ORF72*-ALS and controls

Comparison of the gene expression profiles of *C9ORF72*-ALS patients (n=10), non-*C9ORF72*-ALS patients (n=16) and controls (n=10) with a threshold of pLikeValue <0.05 identified 319 differentially expressed probe sets between *C9ORF72*-ALS patients and controls, 150 differentially expressed probe sets between non-*C9ORF72*-ALS patients and controls, and 294 differentially expressed probe sets between *C9ORF72*-ALS patients and non-*C9ORF72*-ALS patients. The top differentially expressed gene, as determined by fold change, between *C9ORF72*-ALS patients and controls was CXCL10 (fold change −1.9), which was the second highest differentially

Table 2 Summary of neuroinvestigative findings in patients with MS-ALS

Patients	MRI findings	EMG findings	CSF oligoclonal bands
Patient 1	Multiple periventricular and infratentorial lesions. No Gd-enhancing lesions	Fibrillation and fasciculation potentials in the right vastus medialis, 1st dorsal interosseous, biceps brachii, rectus abdominis, masseter and genioglossus. Increased amplitude and duration of motor units+reduced interference pattern in the right biceps brachii, rectus abdominis and masseter	+
Patient 2	Multiple lesions in the corpus callosum, cerebellum and supratentorial lesions. High signal lesion in the right posterior aspect of cervical spinal cord (C4). No Gd-enhancing lesions	Fibrillation and fasciculation potentials and positive sharp waves in the right biceps, 1st dorsal interosseous, tibialis anterior and left triceps, gastrocnemius and quadriceps. Increased amplitude and duration of motor units+reduced interference pattern in the right biceps, left gastrocnemius, triceps and quadriceps	+
Patient 3	Multiple periventricular and juxtacortical lesions. Multiple lesions in the cervical spinal cord (C2, C4–C5) and a single lesion at the dorsal spinal cord (T2). No Gd-enhancing lesions	Fibrillations, fasciculations and positive sharp waves in the right deltoid, abductor pollicis brevis, tibialis anterior and vastus medialis. Increased amplitude and duration of motor unit potentials +reduced interference pattern in the right deltoid, abductor pollicis brevis, tibialis anterior, vastus medialis, genioglossus and thoracic paraspinals	+
Patient 4	Multiple lesions in the periventricular white matter, pontine base and cerebellar hemispheres. No Gd-enhancing lesions	Fibrillations and positive sharp waves in the right quadriceps, tibialis anterior and gastrocnemius. Increased amplitude and duration of motor units with unstable components +reduced interference pattern in the right biceps and 1st dorsal interosseous	+
Patient 5	Multiple lesions within the periventricular white matter, subcortical white matter, upper cervical cord, brainstem, cerebellum and corpus callosum. Several of the lesions showed gadolinium enhancement indicative of active disease	Fibrillation and fasciculation potentials and positive sharp waves in the right biceps brachii, abductor pollicis brevis, extensor digitorum communis, tibialis anterior, vastus lateralis, genioglossus and thoracic paraspinals. Increased amplitude and duration of motor units with unstable components+reduced interference pattern in the right biceps brachii, abductor pollicis brevis, extensor digitorum communis, tibialis anterior, vastus lateralis and thoracic paraspinals	+
Patient 6	Multiple lesions within the periventricular white matter. No Gd-enhancing lesions	Fibrillation and fasciculation potentials in the left deltoid, extensor digitorum communis, vastus medialis and tibialis anterior. Increased amplitude and duration of motor units+reduced interference pattern in the left deltoid, extensor digitorum communis, vastus medialis and tibialis anterior	+
Patient 7	Multiple lesions within the periventricular white matter and corpus callosum. No Gd-enhancing lesions	Fibrillation and fasciculation potentials in the right biceps brachii, extensor digitorum communis, 1st dorsal interosseous, vastus medialis, tibilais anterior and genioglossus. Increased amplitude and duration of motor units+reduced interference pattern in the right biceps brachii, extensor digitorum communis, 1st dorsal interosseous, vastus medialis and tibialis anterior	–

CSF, cerebrospinal fluid; EMG, electromyography, Gd, gadolinium, MS-ALS, multiple sclerosis-amyotrophic lateral sclerosis; +, present; –, absent.

expressed gene between *C9ORF72*-ALS patients and non-*C9ORF72*-ALS patients (fold change −1.4). Analysis with MetaCore to identify known transcription factor targets enriched in the differentially expressed genes between *C9ORF72*-ALS patients and controls and between non-*C9ORF72*-ALS patients and controls highlighted the NF-κB pathway in both ($p<0.000001$).

Results of quantitative polymerase chain reaction for CXCL10 mRNA in lymphoblastoid cells from *C9ORF72*-ALS, non-*C9ORF72*-ALS and controls

QPCR comparison of *CXCL10* expression in lymphoblastoid cells between *C9ORF72*-ALS and control samples and between *C9ORF72*-ALS and non-*C9ORF72*-ALS samples confirmed the findings of the microarray in the same patients. CXCL10 was significantly reduced in C9ORF72-ALS cases compared with controls (fold change −1.7, $p<0.05$) and in C9ORF72-ALS cases compared with non-C9ORF72-ALS cases (fold change −1.7, $p<0.001$; figure 3).

Results of multiplexed fluorescent bead-based immunoassay of CSF from *C9ORF72*-ALS, non-*C9ORF72*-ALS and controls

Measurement of CXCL10 concentrations in CSF from *C9ORF72*-ALS cases (n=4), non-*C9ORF72*-ALS cases (n=5) and controls (n=5) confirmed the findings of the transcriptome analysis at a protein level in a different group of patients. The CXCL10 concentration was highest in non-*C9ORF72*-ALS patients (mean 80.7 pg/ml, SD 31.8), lower in controls (mean

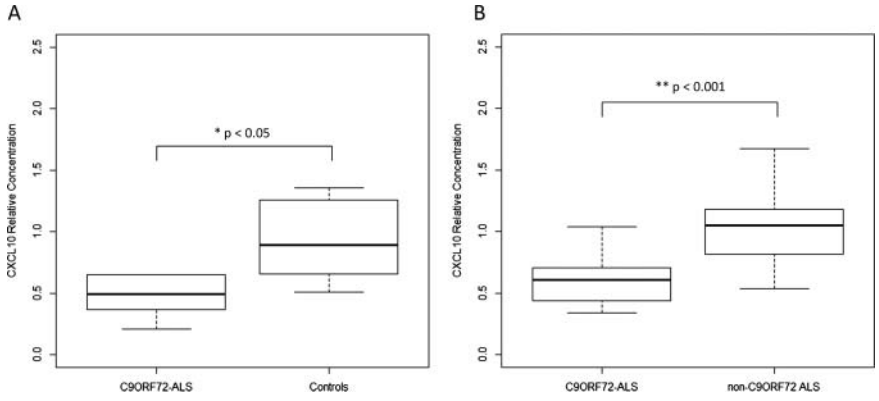


Figure 3 Quantitative polymerase chain reaction measurement of CXCL10 mRNA in lymphoblastoid cells from C9ORF72-ALS patients, non-C9ORF72 ALS patients and controls. Levels of *CXCL10* expression were reduced in C9ORF72-ALS patients compared with controls (A). Levels of *CXCL10* expression were reduced in C9ORF72-ALS patients compared with non-C9ORF72 ALS patients (B). ALS, amyotrophic lateral sclerosis.

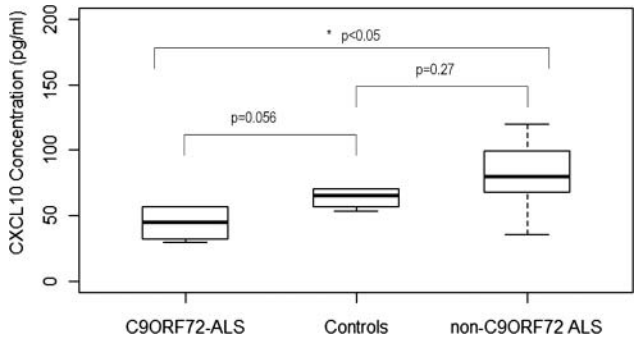


Figure 4 Measurement of CXCL10 concentration in cerebrospinal fluid. Levels of CXCL10 were higher in non-C9ORF72 ALS patients than in controls and C9ORF72-ALS patients. The lowest levels were found in C9ORF72-ALS patients. ALS, amyotrophic lateral sclerosis.

72.1 pg/ml, SD 24.4) and lowest in *C9ORF72*-ALS patients (mean 44.5 pg/ml, SD 14.3) (figure 4). The difference between *C9ORF72*-ALS and non-*C9ORF72*-ALS patients reached statistical significance ($p<0.05$).

DISCUSSION

We and others have shown an association between MS and ALS. We have demonstrated that it is unlikely that this is a chance occurrence, but definitive proof of this rests on the demonstration of a mechanism of association. A genetic basis is suggested by the observation that families with MS have a higher risk of ALS.[22 23] We have highlighted two possible mechanisms based on this premise and the findings in our cases.

Association based on genetic risk factors for autoimmunity

To date, the strongest genetic risk factors for MS involve the human leucocyte antigen (HLA)-associated genes. There is evidence that our patients with MS-ALS share genetic risk factors for autoimmunity: patient 4 had subclinical hyperthyroidism and a strong family history of thyroid disease; patient 7 had hypothyroidism; patients 1 and 6 had a family history of MS. A similar association was previously described in MS-ALS by Dynes *et al*.[5] An association between ALS and HLA genotypes is controversial. In the UK, early efforts to determine the association of ALS and HLA genes revealed an increased incidence of HLA antigens A2 and A28 in ALS in Scotland[24] and a reduced incidence of the DR4 antigen in ALS in northern England.[25] Since then, inconsistent findings have arisen from various centres globally. Given our findings, it appears likely that a genetic susceptibility for ALS in MS is related to hexanucleotide expansions of *C9ORF72,* rather than a predisposition to autoimmune disease.

Association based on expansion of *C9ORF72*

It has been suggested that the penetrance of *C9ORF72* expansions is relatively low, based on the relative abundance of apparently sporadic *C9ORF72*-ALS compared with *C9ORF72*-ALS cases with clear familial disease. In our population, the *C9ORF72* expansion was present in 0.6% of controls.[11] Given that our screen of 215 MS patients revealed no cases with a *C9ORF72* expansion, it seems likely that *C9ORF72* is not commonly associated with MS per se. This is supported by the observation that the daughter of patient 1 has developed MS although she does not have a *C9ORF72* expansion, which is in keeping with a cause of MS independent of the repeat sequence in both the daughter and patient 1 (figure 5). Alternatively, MS in a patient with a *C9ORF72* expansion may increase the likelihood of this genetic change becoming penetrant and causing ALS. This raises the possibility that MS modifies pathways relevant to *C9ORF72*-ALS. Consistent with this, the *C9ORF72*-ALS disease process appears to be more rapid in the presence of concurrent MS than in pure *C9ORF72*-ALS. In a similar manner, oligogenic variants of familial ALS have been reported with expansions of *C9ORF72* in combination with other pathogenic mutations,[26] thus perhaps an additional modifier is a requirement for penetrance more broadly in *C9ORF72*-ALS.

There is evidence for similar patterns of neuroinflammation in MS and ALS.[27] More specifically, inflammatory pathways

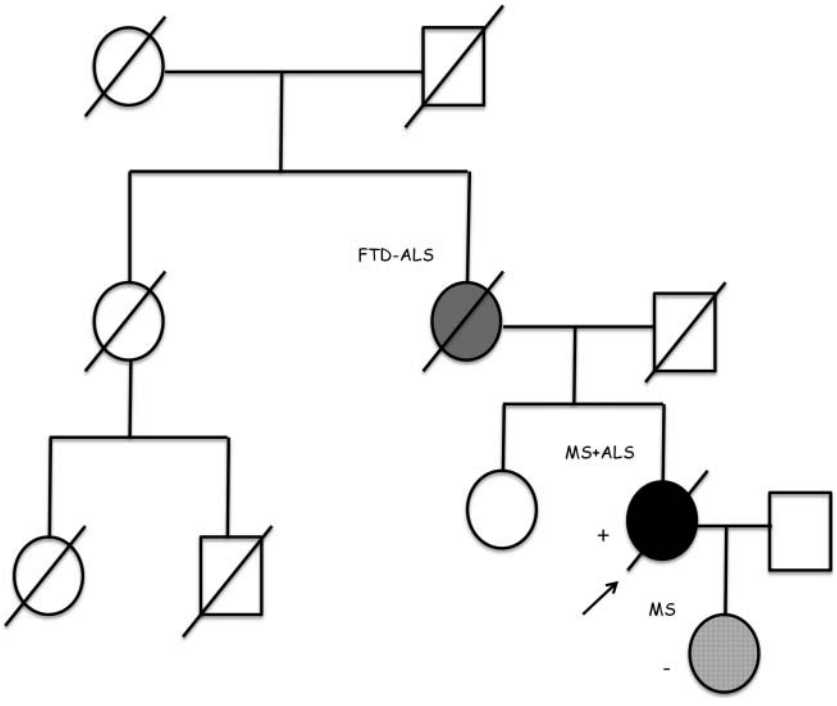


Figure 5 Family tree for patient 1. The index case with multiple sclerosis (MS) and amyotrophic lateral sclerosis (ALS) is indicated by the arrow. The *C9ORF72* expansion was present in the patient with MS-ALS but absent from her daughter who suffered only MS. +=carrier of expansion; −=confirmed absence of the expansion.

related to the transcription factor NF-κB have been associated with MS and ALS. Activated NF-κB is increased in the nuclei of oligodendrocytes in active MS plaques but not within inactive plaques,[28] and a polymorphism in *NFKB1* which encodes a subunit of NF-κB has been associated with MS.[29] Equally, NF-κB activation has been demonstrated in astrocytes from patients with ALS,[12] and mutations of *OPTN* implicated in ALS abolish the ability of optineurin to block NF-κB activation.[30] Furthermore, inhibition of NF-κB activity ameliorated the disease phenotype in transgenic mice expressing a mutated form of TDP-43 implicated in ALS.[31] Data presented here from gene expression profiling suggest that a significant number of genes controlled by NF-κB are differentially expressed in all forms of ALS. One cytokine known to be under the control of NF-κB is CXCL10, which is reported to be elevated in CSF from patients with MS.[32] In CSF from patients with ALS, CXCL10 is elevated and levels of CXCL10 inversely correlate with the rate of disease progression, suggesting a neuroprotective effect.[33] We have shown that, in *C9ORF72*-ALS patients but not in non-*C9ORF72*-ALS patients, CXCL10 is downregulated at the transcriptome level in lymphocytes and at the protein level in CSF. We postulate that the absence of this neuroprotective effect may contribute to the observed predisposition of *C9ORF72* expansion carriers to develop ALS in the context of MS and NF-κB activation. This finding will need confirmation and developing further in a larger series of samples including patients with MS-ALS.

CONCLUDING REMARKS

The coexistence of MS and ALS is rare. However, our case series highlights that more than 1% of patients with ALS may have a preceding history of MS, suggesting more than a chance association. Clinicians should be aware of the possible development of ALS in the face of unusually rapid progression of disability and the presence of lower motor neuron signs in patients with an established diagnosis of MS.

In addition, we have presented evidence that a large proportion of such MS-ALS cases are associated with a hexanucleotide repeat expansion of *C9ORF72*, and that MS is a poor prognostic marker for *C9ORF72*-ALS but not for ALS more broadly. We have suggested that the development of MS is a trigger which facilitates the penetrance of the *C9ORF72* expansion, perhaps through pathways related to NF-κB. This is an exciting step towards identification of a therapeutic target in these patients and also more broadly in ALS and MS.

Acknowledgements We are very grateful to all of the patients with ALS and MS who donated biosamples.

Contributors JCK and AI are joint first authors, BS and PJS are joint senior authors. JCK, AI, JRH, JK, BS and PJS contributed to the writing and editing of the paper. JCK, SP, PJS, BS, JL, CC, CMD, SS and DAS contributed to the identification of cases and obtaining DNA samples. JCK and AM carried out the screening for the C9ORF72 expansion. SJW helped with the statistical analysis.

Funding The biosample collection for the ALS cases was supported by the MND Association and the Wellcome Trust (PJS). Work on MS samples was supported by the Cambridge NIHR Biomedical Research Centre. Genetic testing was supported by NIHR CLAHRC for South Yorkshire. PJS and JK are supported by an EU Framework 7 grant (Euromotor No 259867). JRH is supported by an MND Association /Medical Research Council Lady Edith Wolfson Fellowship award (G0 800380).

Competing interests None.

Ethics approval Ethics approval was obtained from South Sheffield Research Ethics Committee and Thames Valley Multi-Centre Research Ethics Committee.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Within SITraN we operate a data management and sharing plan fully aligned with the Wellcome Trust and MRC policy and requirements for public access to data derived from grant funding. Metadata, curation and data preservation are enhanced through our access to expert biostatistical support (SITraN Computational Biology Group). We propose to operate a policy of time-limited exclusivity in relation to: primary analyses; publications, patents and other outputs. This embargo on public access will be restricted to 3 years downstream of the initial discovery. Policy for sharing data with others, including the management and prioritisation of access to data: As above, all data will be made publically available subject to time-limited exclusivity for the purposes specified. We encourage, but do not demand, collaborative interaction with groups who access the data to ensure that they fully understand the data, its sources and limitations, and its responsible use. No data that could undermine patient confidentiality will be released to public archiving. We do not foresee any issues related to the prioritisation of the release of datasets.

REFERENCES

1. **Logroscino G,** Traynor BJ, Hardiman O, *et al*. Incidence of amyotrophic lateral sclerosis in Europe. *J Neurol Neurosurg Psychiatry* 2010;**81**:385–90.
2. **Compston A,** Coles A. Multiple sclerosis. *Lancet* 2008;**372**:1502–17.
3. **Allen JA,** Stein R, Baker RA, *et al*. Muscle atrophy associated with multiple sclerosis: a benign condition or the onset of amyotrophic lateral sclerosis? *J Clin Neurosci* 2008;**15**:706–8.
4. **Confavreux C,** Moreau T, Jouvet A, *et al*. Association of amyotrophic lateral sclerosis and multiple sclerosis. *Rev Neurol (Paris)* 1993;**149**:351–3.
5. **Dynes G,** Schwimer C, Staugaitis S, *et al*. Amyotrophic lateral sclerosis with multiple sclerosis: a clinical and pathological report. *Amyotroph Lateral Scler Other Motor Neuron Disord* 2000;**1**:349–53.
6. **Li G,** Esiri M, Ansorge O, *et al*. Concurrent multiple sclerosis and amyotrophic lateral sclerosis: where inflammation and neurodegeneration meet? *J Neuroinflammation* 2012;**9**:20.
7. **Machner B,** Gottschalk S, Kimmig H, *et al*. Kombiniertes Auftreten von amyotropher Lateralsklerose und Multipler Sklerose. *Der Nervenarzt* 2007;**78**:1440–3.
8. **Hader W,** Rpzdilsky B, Nair C. The concurrence of multiple sclerosis and amyotrophic lateral sclerosis. *Can J Neurol Sci* 1986;**13**:66–9.
9. **Renton AE,** Majounie E, Waite A, *et al*. A Hexanucleotide repeat expansion in C9ORF72 Is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 2011;**72**:257–68.
10. **DeJesus-Hernandez M,** Mackenzie IR, Boeve BF, *et al*. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 2011;**72**:245–56.
11. **Cooper-Knock J,** Hewitt C, Highley JR, *et al*. Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. *Brain* 2012;**135**:751–64.
12. **Migheli A,** Piva R, Atzori C, *et al*. c-Jun, JNK/SAPK kinases and transcription factor NF-kappa B are selectively activated in astrocytes, but not motor neurons, in amyotrophic lateral sclerosis. *J Neuropathol Exp Neurol* 1997;**56**:1314–22.
13. **Yan J,** Greer J. NF-kappa B, a potential therapeutic target for the treatment of multiple sclerosis. *CNS Neurol Disord Drug Targets* 2008;**7**:536–57.
14. **Polman CH,** Reingold SC, Banwell B, *et al*. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol* 2011;**69**:292–302.
15. **Brooks B.** El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/ Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial "Clinical limits of amyotrophic lateral sclerosis" workshop contributors. *J Neurol Sci* 1994;**124**:96–107.
16. **Ferraiuolo L,** Heath PR, Holden H, *et al*. Microarray analysis of the cellular pathways involved in the adaptation to and progression of motor neuron injury in the SOD1 G93A mouse model of familial ALS. *J Neurosci* 2007;**27**:9201–19.
17. **Pearson R,** Liu X, Sanguinetti G, *et al*. puma: a Bioconductor package for propagating uncertainty in microarray analysis. *BMC Bioinformatics* 2009;**10**:211.
18. **Liu X,** Milo M, Lawrence ND, *et al*. Probe-level measurement error improves accuracy in detecting differential gene expression. *Bioinformatics* 2006;**22**:2107–13.
19. **Kirby J,** Halligan E, Baptista MJ, *et al*. Mutant SOD1 alters the motor neuronal transcriptome: implications for familial ALS. *Brain* 2005;**128**:1686–706.
20. **Shepherd D,** Summers A. Prevalence of multiple sclerosis in Rochdale. *J Neurol Neurosurg Psychiatry* 1996;**61**:415–17.
21. **Confavreux C,** Aimard G, Devic M. Course and prognosis of multiple sclerosis assessed by the computerized data processing of 349 patients. *Brain* 1980;**103**:281–300.
22. **Hemminki K,** Li X, Sundquist J, *et al*. Risk for multiple sclerosis in relatives and spouses of patients diagnosed with autoimmune and related conditions. *Neurogenetics* 2009;**10**:5–11.
23. **Hemminki K,** Li X, Sundquist J, *et al*. Familial risks for amyotrophic lateral sclerosis and autoimmune diseases. *Neurogenetics* 2009;**10**:111–16.

24. **Behan PO,** Dick HM, Durward WF. Histocompatibility antigens associated with motor neurone disease. *J Neurol Sci* 1977;**32**:213–17.
25. **Woo E,** Nightingale S, Dick D, *et al*. A study of histocompatibility antigens in patients with motor neuron disease in the northern region of England. *J Neurol Neurosurg Psychiatry* 1986;**49**:435–7.
26. **van Blitterswijk M,** van Es MA, Hennekam EAM, *et al*. Evidence for an oligogenic basis of amyotrophic lateral sclerosis. *Hum Mol Genet* 2012;**21**:3776–84.
27. **Rentzos M,** Rombos A, Nikolaou C, *et al*. Interleukin-15 and interleukin-12 are elevated in serum and cerebrospinal fluid of patients with amyotrophic lateral sclerosis. *Eur Neurol* 2010;**63**:285–90.
28. **Bonetti B,** Stegagno C, Cannella B, *et al*. Activation of NF-κB and c-jun transcription factors in multiple sclerosis lesions: implications for oligodendrocyte pathology. *Am J Pathol* 1999;**155**:1433–38.
29. **Zhang R,** Hadlock KG, Do H, *et al*. Gene expression profiling in peripheral blood mononuclear cells from patients with sporadic amyotrophic lateral sclerosis (sALS). *J Neuroimmunol* 2011;**230**:114–23.
30. **Maruyama H,** Morino H, Ito H, *et al*. Mutations of optineurin in amyotrophic lateral sclerosis. *Nature* 2010;**465**:223–6.
31. **Swarup V,** Phaneuf D, Dupré N, *et al*. Deregulation of TDP-43 in amyotrophic lateral sclerosis triggers nuclear factor κB–mediated pathogenic pathways. *J Exp Med* 2011;**208**:2429–47.
32. **Sørensen TL,** Sellebjerg F, Jensen CV, *et al*. Chemokines CXCL10 and CCL2: Differential involvement in intrathecal inflammation in multiple sclerosis. *Eur J Neurol* 2001;**8**:665–72.
33. **Tateishi T,** Yamasaki R, Tanaka M, *et al*. CSF chemokine alterations related to the clinical course of amyotrophic lateral sclerosis. *J Neuroimmunol* 2010;**222**:76–81.

Concurrence of multiple sclerosis and amyotrophic lateral sclerosis in patients with hexanucleotide repeat expansions of *C9ORF72*

Azza Ismail, Johnathan Cooper-Knock, J Robin Highley, Antonio Milano, Janine Kirby, Emily Goodall, James Lowe, Ian Scott, Cris S Constantinescu, Stephen J Walters, Sian Price, Christopher J McDermott, Stephen Sawcer, D Alastair S Compston, Basil Sharrack and Pamela J Shaw

J Neurol Neurosurg Psychiatry 2013 84: 79-87 originally published online October 20, 2012
doi: 10.1136/jnnp-2012-303326

3. Molecular pathogenesis of *C9orf72*-disease

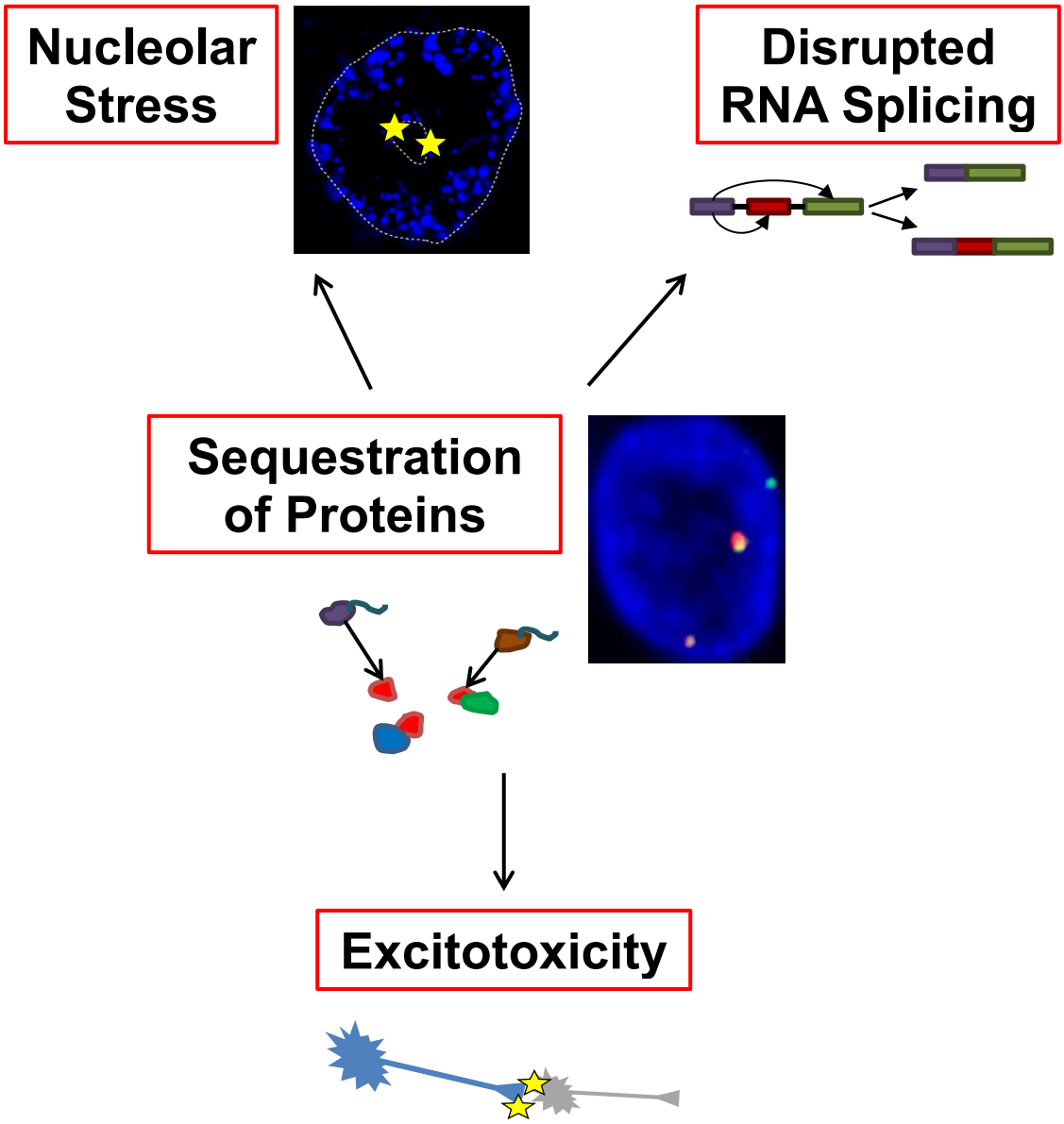
The mechanism of neurotoxicity in *C9orf72*-disease is unknown. However evidence is being gathered for three potential mechanisms, all of which have precedence in other neurological diseases mediated by repeat expansions: 1) RNA based toxicity of the transcribed repeat; 2) protein based toxicity via translation of the expanded RNA to form dipeptide repeat proteins (DPRs) and 3) haploinsufficiency.

3.1. RNA toxicity

RNA foci formed from the repeat sequence were observed in the earliest studies of *C9orf72*-disease (DeJesus-Hernandez *et al.* 2011). The numbers of RNA foci have been correlated with pathogenic severity in cell models (Donnelly *et al.* 2013, Sareen *et al.* 2013), and in tissue from FTD cases (Lagier-Tourenne *et al.* 2013). A number of molecular phenotypes have been linked to the RNA foci (Donnelly *et al.* 2013, Lagier-Tourenne *et al.* 2013, Lee *et al.* 2013, Mizielinska *et al.* 2013, Sareen *et al.* 2013, Haeusler *et al.* 2014). Most groups appear to be exploring the idea that RNA foci sequester and therefore alter the function of certain proteins, including for example ADARB2 leading to excitotoxicity (Donnelly *et al.* 2013) and nucleolin leading to nucleolar stress (Haeusler *et al.* 2014). These mechanisms are summarised in **Figure 2**. It is now clear that RNA foci are formed from transcription of the GGGGCC-repeat in the sense and the antisense direction (Gendron *et al.* 2013, Lagier-Tourenne *et al.* 2013, Mizielinska *et al.* 2013, Haeusler *et al.* 2014).

In **paper 5** we have examined the localisation and interactions of sense RNA foci (Cooper-Knock *et al.* 2014). We performed *in vitro* RNA pulldown using biotinylated $(GGGGCC)_5$ molecules. This RNA was mixed with whole and nuclear homogenates from SH-SY5Y cells, and with homogenate of cerebellar granule neurons. Granule

Figure 2 (reproduced from Cooper-Knock, Kirby et al): Proposed mechanisms of toxicity mediated by protein sequestration to RNA foci. Various proteins have been shown to be sequestered to RNA foci transcribed either in the sense or antisense direction from the GGGGCC-repeat sequence. Proposed mechanisms include sequestration of proteins important to mRNA-splicing with consequent disruption of RNA splicing; sequestration of ADARB2 leading to increased susceptibility to excitotoxicity; and sequestration of nucleolin producing nucleolar stress.



neurons were chosen because they represent a relatively easy to access neuronal population which is specifically targeted by C9orf72-disease (see section *3.2*). Binding partners were identified by mass spectroscopy. By collating the different lists 103 unique binding partners were identified; these binding partners were significantly enriched for RNA recognition motif (RRM) containing proteins, many of which are key components of nuclear speckles which are nuclear domains thought to function in the storage and supply of splicing factors for transcription (Spector *et al.* 2011). We selected certain of these for further validation. Firstly we performed UV-crosslinking to demonstrate direct and specific interaction with our biotinylated RNA molecule. Secondly immunohistochemistry (IHC) in post mortem tissue from *C9orf72*-ALS patients, non-*C9orf72* ALS patients and controls demonstrated colocalisation of hnRNP A1, hnRNP H1/F, ALYREF and SRSF2 with sense RNA foci in CNS neurons including motor neurons of the ventral horn. RNA foci were not seen in non-*C9orf72* cases. Colocalisation of each protein was observed with less than a third of RNA foci; we interpreted this as suggestive of dynamic sequestration. In view of the relatively late age of onset of *C9orf72*-disease and the extremely variable phenotype, we propose a model whereby dynamic sequestration of a relatively large number of RRM-containing proteins might have a low-level effect on nuclear speckle function, which in time might precipitate disease (Cooper-Knock *et al.* 2014). Interestingly we and others have observed RNA foci in tissue from pre-symptomatic patients with no clinical disease (**paper 5**, (Lagier-Tourenne *et al.* 2013, Cooper-Knock *et al.* 2014)).

In **paper 6** we investigated the relative importance of sense and antisense RNA foci to the pathophysiology of neuronal injury. We demonstrated that the expression patterns of sense and antisense RNA foci are distinct; sense RNA foci are present at

higher frequency in granule neurons of the cerebellum whilst antisense RNA foci are present at higher frequency in Purkinje neurons of the cerebellum and in motor neurons of the ventral horn. Consistent with these findings, we have shown that DPRs derived from antisense RNA are present at higher frequency in motor neurons and DPR derived from sense RNA are present at higher frequency in cerebellar granule neurons. All of these populations have been observed to show neurodegeneration in *C9ORF72*-disease, but motor neurons are the primary target of pathology in ALS. Also in **paper 6** we examined interaction between antisense RNA foci and selected RRM-containing proteins including SRSF2, hnRNP K, hnRNP A1, ALYREF and hnRNP H/F. IHC demonstrated colocalisation at comparable frequencies to those observed with sense RNA foci in **paper 5.** Direct interactions were confirmed by UV-crosslinking using a biotinylated $(CCCCGG)_5$ RNA. Our data confirmed the findings of others that binding partners of the sense and antisense RNA foci are broadly similar (Haeusler *et al.* 2014). Therefore, if their interactions are key to *C9ORF72*-pathogenesis, then both species should equally toxic.

Finally, in **paper 6** we showed that the presence of antisense RNA foci, but not sense RNA foci, is significantly associated with nuclear clearance of TDP-43 in motor neurons, a hallmark of ALS pathology. Since their interactions are similar, we suggest that the increased frequency of antisense RNA foci in motor neurons is key to this observation – the antisense foci might be expected to overwhelm the effect of the less numerous sense foci. This has implications for development of therapies which must target both forms of RNA foci.

3.1.1 The effect of C9orf72 expansions on the transcriptome

Since we and others have observed interactions between RNA foci and RRM-containing proteins, in **paper 7** we decided to determine the effect of the expansion

on the transcriptome. We utilised RNA extracted from laser captured motor neurons and from lymphoblastoid cell lines derived from C9orf72-ALS patients. Gene-level analysis in both cell types demonstrated up-regulation of genes encoding RRM-containing proteins, which may represent compensation for sequestration and is consistent with our model. Importantly this helped to validate our use of the lymphoblastoid cell lines, with better quality RNA and larger sample numbers, in the analysis of exon level splicing. In view of our proposal of a relatively low-level effect on nuclear speckle function we decided to take a novel approach. We analysed consistency of splicing events within sample groups as a proxy for the error rate in splicing. Splicing errors are a normal component of biological function and are thought to play a role in establishing necessary genetic diversity (Pickrell et al. 2010); however it is conceivable that excessive errors might be pathogenic. We described an increase in the splicing error rate in lymphoblastoid cell lines derived from C9orf72-ALS patients, which is more pronounced in samples derived from patients with rapidly progressive disease. Consistency has been missing from studies of the transcriptome in C9orf72-disease and our data might explain this. Moreover, the frequency of RNA foci within lymphoblasotid cells was higher in patients with a more rapid disease course, although this did not reach statistical significance.

3.2. Protein toxicity

Precedence from other repeat expansion disorders led to the search for and the discovery of DPRs in *C9orf72*-disease (Mori *et al.* 2013). Antibodies were developed to the five different proteins corresponding to the six possible reading frames for the sense (poly-GA, poly-GR, poly-GP) and antisense (poly-PA, poly-PR, poly-GP)

repeat-RNAs: all of the protein species have been identified within ubiquitinated NCIs and in several cases were observed to co-aggregate (Mori *et al.* 2013).

The poly-PR and poly-GP proteins are associated with a potential ATG start codon, but the other species must be translated via a non-canonical mechanism which is as yet unknown and may mediate translation of all of the proteins. It is proposed that this mechanism is translation initiated directly by the repeat expansion, as has previously been observed (Zu *et al.* 2011), so called 'repeat associated, non-ATG' (RAN) translation. In this study of spinocerebellar ataxia 8, a neuromuscular disorder caused by a CAG-repeat expansion of the ataxin 8 gene, Zu *et al* showed that translation of the repeat sequence occurred independently of the presence of an ATG-site. As with the *C9orf72*-expansion, translation was demonstrated in three independent reading frames corresponding to poly-G, poly-S and poly-A. Interestingly, the relative mix of these protein species varied depending on the repeat length. They also concluded that hairpin secondary structure in the RNA is important for RAN-translation since a CAA-repeat with similar properties, but without the ability to form a hairpin secondary structure, was translated only in the presence of an ATG start codon. This work has already had significant effect on the study of *C9orf72*-disease and it is likely that it will continue to do so; interestingly both antisense and sense RNA molecules are thought to able to form a hairpin secondary structure (Haeusler *et al.* 2014). Work with a transfected GGGGCC-repeat sequence has demonstrated length dependence of translation of the poly-PR and poly-GP proteins in a cell model (Gendron *et al.* 2013).

DPR proteins appear to be toxic in cell and animal models (Zu *et al.* 2013, Kwon *et al.* 2014, Mizielinska *et al.* 2014). However, neuropathological studies find no relationship between the extent of the DPR pathology and clinical severity (Davidson

et al. 2014). In fact, levels of the aberrantly translated protein appear to inversely correlate with vulnerability of different neuronal groups to neurodegeneration in autopsy material, in direct contrast to the levels of TDP-43-positive inclusions (Millecamps *et al.* 2012). This may be consistent with a protective role for the formation of DPRs. However, it should be noted that this was a study of pathological material and at the end-stage of disease it is impossible to rule out the possibility that the neurons which have already died are the ones containing the highest burden of DPRs. Time course studies in model systems will be required to shed light on this ambiguity. Recently, three studies have moved this story forwards. Kwon *et al* have provided evidence from a cell model that the poly-GR and poly-PR DPRs may bind irreversibly to the nucleolus, leading to toxicity via disruption of pre-mRNA splicing and ribosome synthesis (Kwon *et al.* 2014). The association of SR domains with nucleoli has been observed previously and shown to be dependent on phosphorylation (Bubulya *et al.* 2004). Kwon *et al* have extended this finding to SR-domain containing proteins implicated in ALS. Moreover they suggest that the poly-GR and poly-PR proteins associate irreversibly with the nucleolus because, unlike SR-domains, they lack serine residues and are therefore unable to undergo phosphorylation by CLK1/2 protein kinases. The nucleolar target of arginine-rich domains is unknown, but the authors speculate that the irreversible binding of DPRs might disrupt the normal processing, and therefore the function, of SR proteins in pre-mRNA splicing, and might also disrupt the normal function of the nucleolus in the synthesis of ribosomal RNA. As evidence in support of this hypothesis, they demonstrate changes in the splicing of the excitatory amino acid transporter-2 (EAAT2) and in the production of small nucleolar RNA (snoRNA) upon administration of PR_{20}. snoRNAs are regulatory molecules important for the synthesis of other

RNAs including ribosomal RNAs. Although narrow, these changes are reminiscent of molecular phenotypes reported previously in ALS (Lin *et al.* 1998). Intriguingly this offers an alternative mechanism for the nucleolar stress observed by Haeusler *et al* described above (Haeusler *et al.* 2014).

This proposed model was strengthened shortly after by the work of Mizielinska *et al* who demonstrated neurodegeneration in a drosophila model upon expression of poly-GP and poly-PR proteins. Toxicity was not present on expression of poly-PA protein and was relatively minor upon expression of poly-GA protein (Mizielinska *et al.* 2014). Importantly, this study included an 'RNA-only' model by expression of stop-codon interrupted GGGGCC repeats which produced RNA foci but not DPRs. The 'RNA-only' model formed RNA foci but was not toxic at any expansion length tested which included 36, 108 and ~288 repeats. The authors note that the repeat lengths tested are much shorter than expansion lengths present in *C9orf72*-disease patients and therefore this result does not exclude RNA toxicity. They did not describe any length dependence of the DPR toxicity, but they did show amelioration of toxicity by inhibition of protein translation.

Finally, May *et al* examined the pattern of expression of the five dipeptide species in HEK293 cells (May *et al.* 2014). The arginine-rich peptides produced dot-like intranuclear inclusions which may correspond to the nucleolar sequestration proposed by Kwon *et al* (Kwon *et al.* 2014). May *et al* described a potential mechanism of pathogenesis associated with poly-GA repeats. They reported ubiquitination and toxicity of cytoplasmic poly-GA inclusions associated with interaction between poly-GA and a number of components of the ubiquitin-proteasome system including Unc119. Unc119 colocalised with poly-GA inclusions

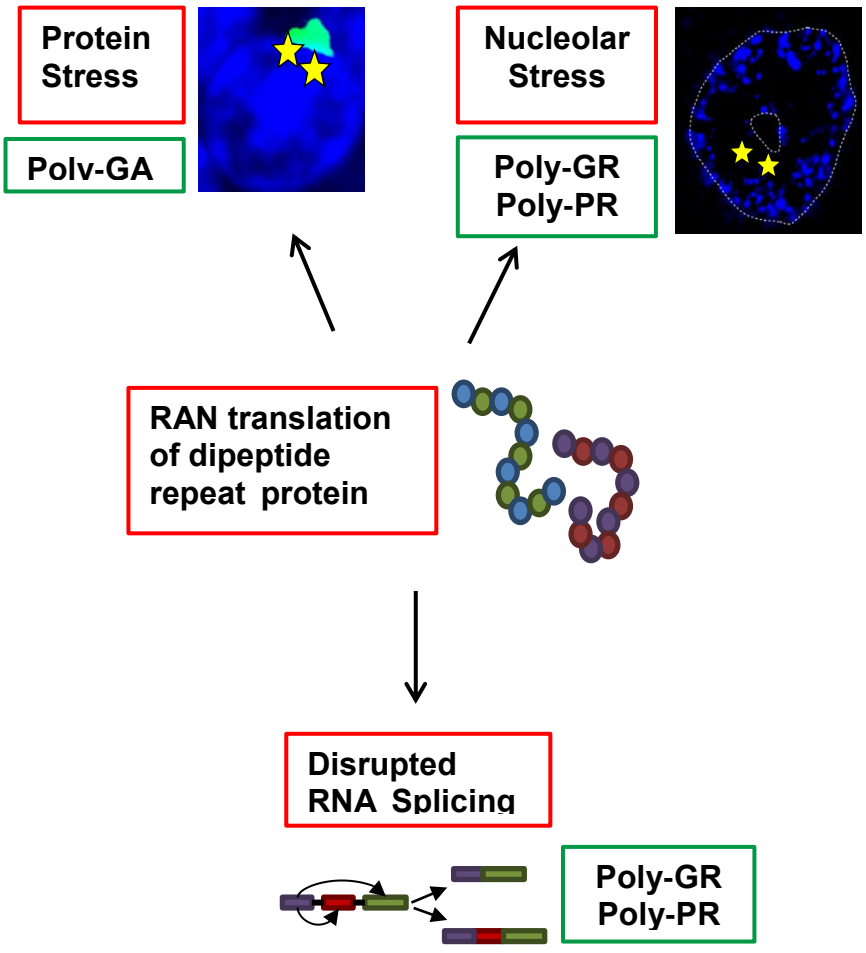
in brains of patients with *C9orf72*-FTD and knockdown of unc119 was neurotoxic in a cell model.

This interesting set of observations, summarised in **Figure 3**, awaits confirmation and further development. In particular, the effect of dipeptide-repeat length is unknown and a direct link to the disease remains to be established.

Our observations in **paper 6** inform this picture: we have shown that spinal motor neurons from *C9orf72*-ALS patients preferentially express antisense foci and antisense RNA derived DPRs. This suggests that the observations of May et al regarding poly-GA inclusions may be less important in this neuronal population. Moreover, our data suggest that of the arginine rich DPRs, poly-PR is likely to be more important to the development of an ALS phenotype than poly-GR which is derived from sense RNA.

Observations of RNA foci have a bearing on the formation of DPRs since inappropriate nuclear export is required to allow repeat-RNA molecules to access the translation machinery. In **paper 5** and **paper 6** we observed cytoplasmic RNA foci in CNS neuronal populations, including motor neurons, as have others (Mizielinska *et al.* 2013, Cooper-Knock *et al.* 2014). In a post-mitotic neuron this could not be achieved by nuclear extrusion during mitosis. We also showed in **paper 5** and **paper 6** that sense and antisense RNA foci interact directly with export adaptors including ALYREF (Cooper-Knock *et al.* 2014) and we propose that this interaction might inappropriately licence the repeat RNA for export to the cytoplasm. If DPRs are toxic, this represents an attractive therapeutic target. It is interesting to note that pathological analysis suggests that the coincidence in cells of sense/ antisense RNA foci and DPR inclusions is relatively low (Gendron *et al.* 2013) which may suggest that the formation of transcripts into foci and the cytoplasmic export of transcripts for

Figure 3 (reproduced from Cooper-Knock, Kirby et al): Proposed mechanisms of toxicity mediated via translation of dipeptide repeat protein (DPR). Both sense and antisense, or GGGGCC- and GGCCCC-repeat RNA species are observed to be translated into DPRs. The absence of traditional ATG-start sites has led to the proposal that this occurs via a repeat-associated initial or repeat-associated non-ATG (RAN) translation. The various DPR sequences are proposed to be toxic via a number of mechanisms including; disruption of normal processing of RNA-binding proteins associated with the nucleolus leading to nucleolar stress and/or disruption of normal RNA splicing; or disruption of normal protein processing pathways leading to formation of protein inclusions and toxic protein stress. Implicated DPR species are shown for each mechanism.



translation, if not mutually exclusive, are mediated independently. If one of these processes is identified as toxic and one as protective then modulating this decision point may represent a therapeutic target.

3.3. Haploinsufficiency

Reduced expression of *C9orf72* mRNA has been reported in the presence of the expansion (DeJesus-Hernandez *et al.* 2011). However, this finding is not consistent (Sareen *et al.* 2013). Most recently, study of a newly generated C9orf72 antibody suggests that there is reduced expression of the C9orf72 protein in the frontal cortex, but not in the cerebellum of both *C9orf72*-ALS and *C9orf72*-FTD patients (Waite *et al.* 2014).

In **paper 8** we demonstrated that small expansions of approximately 50 repeats do not reduce *C9orf72* transcription (Cooper-Knock *et al.* 2013) possibly because smaller expansions do not lead to hypermethylation of a CpG island 5' to the repeat sequence in the promoter region (Xi *et al.* 2013, Xi *et al.* 2014). If smaller repeat lengths are pathogenic (Byrne *et al.* 2013, Gomez-Tortosa *et al.* 2013), then we suggest that haploinsufficiency is not the responsible mechanism. A caveat to this comes from our observations in **paper 10**: we describe an ALS patient with an intermediate length *C9orf72* expansion inherited on the background of the 9p21 risk haplotype but without typical *C9orf72* neuropathology including extra-motor p62-positive, TDP-43-negative neuronal inclusions, DPR inclusions and RNA foci (Beer *et al.* 2014). We suggest that this patient may actually represent a case of sporadic ALS and that the repeat length in this case is not sufficient to initiate typical *C9orf72*-disease or to produce haploinsuffciency.

Developing this story further, CpG hypermethylation of the *C9orf72* promoter has been shown to correlate with the burden of neuropathology: the presence of promoter hypermethylation is associated with reduced accumulation of DPRs and RNA foci in the CNS of *C9orf72*-patients (Liu *et al.* 2014). Moreover, the same study showed in lymphoblastoid cell lines derived from *C9orf72*-expansion carriers, that demethylation of the promoter led to increased vulnerability of the cells to oxidative and autophagic stress. While not conclusive, this suggests that reduced expression of expanded *C9orf72* might be protective rather than pathogenic. A number of other mechanisms have been proposed for haploinsufficiency: trimethylation of histones H3 and H4 has been reported in *C9orf72*-expansion carriers and linked to increased binding of these histones to the repeat sequence with consequent reduced expression of *C9orf72* (Belzil *et al.* 2013). Finally, biochemical analysis has suggested that formation of expanded *C9orf72* DNA and RNA into hybrid R-loops may also contribute to abortive transcription (Haeusler *et al.* 2014).

Friedreich's ataxia is a neuromuscular disorder associated with an intronic repeat expansion in the *FXN* gene; in this disease the mutation must be homozygous to be pathogenic and haploinsufficiency has been confirmed at the protein level. Drawing parallels with *C9orf72*, the *FXN* repeat expansion has been associated with epigenetic silencing. However, additional mechanisms have been identified: blockage of transcription-elongation by DNA-repeat secondary structure has been demonstrated (Burnett *et al.* 2006) and the presence of the repeat expansion has been shown to reduce levels of mature *FXN* mRNA via interaction with trans-acting splicing factors (Baralle *et al.* 2008). The latter mechanism might explain some of the controversy in measurement of *C9orf72* mRNA since the diversity of splice variants produced might lead to contrasting results in various qPCR assays

depending on the primers utilised. It remains to be seen whether similar mechanisms are at play in *C9orf72*-disease.

Another observation not consistent with a pathogenic role for haploinsufficiency comes from two patients with expansions of both *C9orf72* loci; one a homozygote (Fratta *et al.* 2013) and the other a compound heterozygote described by our group in **paper 8** (Cooper-Knock *et al.* 2013). Both cases suffered FTD, but neither phenotype was outside the usual phenotypic spectrum. This is not consistent with a pure haploinsufficiency model which would predict disease severity in proportion to the number of involved alleles.

Little is known about the normal role of the C9orf72 protein. The most conserved residues in C9orf72 are distributed throughout the protein, suggesting that it functions as a single block (Levine *et al.* 2013). Five out of six independent structure prediction tools used to examine the primary sequence of C9orf72 suggested a 'differentially expressed in normal and neoplasia' (DENN) functional domain (Levine *et al.* 2013). DENN proteins are Rab-GTP/GDP exchange factors. Moreover, examination of the predicted secondary structure of C9orf72 also revealed significant homology with DENN proteins (Levine *et al.* 2013). This has led to the suggestion that C9orf72 is important in the regulation of Rab activity and thus membrane trafficking, a proposal which is supported by another study showing that the C9orf72 protein co-localises in neurons with Rab proteins and membrane vesicles implicated in autophagy and endocytosis (Farg *et al.* 2014). It remains to be seen whether haploinsufficiency is a disease modifier, but this evaluation may be just around the corner if the newly developed antibodies become widely accepted as sensitive and specific.

List of Papers:

Paper 5:

Cooper-Knock J+, Walsh MJ+, Higginbottom A, Highley JR, Dickman MJ, Edbauer D, Ince PG, Wharton SB, Wilson SA, Kirby J, Hautbergue GM, Shaw PJ. Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. **Brain**. 2014; 137:2040-51

- *Joint first author*
- *Performed all of confocal imaging and analysis*
- *Primary role in drafting manuscript*

Paper 6:

Cooper-Knock J, Higginbottom A, Stopford MJ, Highley JR, Ince PG, Wharton SB, Pickering-Brown S, Kirby J, Hautbergue GM, Shaw PJ Antisense RNA foci in the motor neurons of *C9ORF72*-ALS patients are associated with TDP-43 proteinopathy **Acta Neuropathologica** [under review]

- *First author*
- *Performed all of confocal imaging and analysis*
- *Peformed cloning and UV-crosslinking*
- *Primary role in drafting manuscript*

Paper 7:

Johnathan Cooper-Knock[1], Joanna Bury[1], Paul R Heath[1],Matthew Wyles[1], Adrian Higginbottom[1], Matthew Walsh[1], Cartherine Gelsthorpe[1], J Robin Highley[1], Guillaume Hautbergue[1], Magnus Rattray[2], Janine Kirby[1], Pamela J Shaw *C9ORF72* GGGGCC expanded repeats produce splicing dysregulation which correlates with disease severity in amyotrophic lateral sclerosis **Plos One** [under review]

- *First author*
- *Performed all analysis of transcriptome data*

- Performed all of wet lab work with the exception of microarrays.

- Primary role in literature review and drafting manuscript

doi:10.1093/brain/awu120 Brain 2014: 137; 2040–2051 | 2040

BRAIN
A JOURNAL OF NEUROLOGY

Sequestration of multiple RNA recognition motif-containing proteins by *C9orf72* repeat expansions

Johnathan Cooper-Knock,[1,*] Matthew J. Walsh,[1,*] Adrian Higginbottom,[1] J. Robin Highley,[1] Mark J. Dickman,[2] Dieter Edbauer,[3] Paul G. Ince,[1] Stephen B. Wharton,[1] Stuart A. Wilson,[4] Janine Kirby,[1] Guillaume M. Hautbergue[1] and Pamela J. Shaw[1]

1 Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, 385A Glossop Road, Sheffield S10 2HQ, UK
2 Chemical and Biological Engineering, ChELSI Institute, University of Sheffield, Mappin Street, Sheffield, S1 3JD, UK
3 DZNE–German Centre for Neurodegenerative Diseases and Munich Cluster of Systems Neurology (SyNergy), Munich, Germany
4 Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Sheffield, S10 2TN, UK

*These authors contributed equally to this work

Correspondence to: Professor Dame Pamela J Shaw,
Sheffield Institute for Translational Neuroscience (SITraN),
385A Glossop Road,
Sheffield S10 2HQ, UK
E-mail: pamela.shaw@sheffield.ac.uk

GGGGCC repeat expansions of *C9orf72* represent the most common genetic variant of amyotrophic lateral sclerosis and frontotemporal degeneration, but the mechanism of pathogenesis is unclear. Recent reports have suggested that the transcribed repeat might form toxic RNA foci that sequester various RNA processing proteins. Consensus as to the identity of the binding partners is missing and whole neuronal proteome investigation is needed. Using RNA fluorescence *in situ* hybridization we first identified nuclear and cytoplasmic RNA foci in peripheral and central nervous system biosamples from patients with amyotrophic lateral sclerosis with a repeat expansion of *C9orf72* (*C9orf72+*), but not from those patients without a repeat expansion of *C9orf72* (*C9orf72−*) or control subjects. Moreover, in the cases examined, the distribution of foci-positive neurons correlated with the clinical phenotype (*t*-test $P < 0.05$). As expected, RNA foci are ablated by RNase treatment. Interestingly, we identified foci in fibroblasts from an asymptomatic *C9orf72+* carrier. We next performed pulldown assays, with $GGGGCC_5$, in conjunction with mass spectrometry analysis, to identify candidate binding partners of the GGGGCC repeat expansion. Proteins containing RNA recognition motifs and involved in splicing, messenger RNA nuclear export and/or translation were significantly enriched. Immunohistochemistry in central nervous system tissue from *C9orf72+* patients with amyotrophic lateral sclerosis demonstrated co-localization of RNA foci with SRSF2, hnRNP H1/F, ALYREF and hnRNP A1 in cerebellar granule cells and with SRSF2, hnRNP H1/F and ALYREF in motor neurons, the primary target of pathology in amyotrophic lateral sclerosis. Direct binding of proteins to GGGGCC repeat RNA was confirmed *in vitro* by ultraviolet-crosslinking assays. Co-localization was only detected in a small proportion of RNA foci, suggesting dynamic sequestration rather than irreversible binding. Additional immunohistochemistry demonstrated that neurons with and without RNA foci were equally likely to show nuclear depletion of TDP-43 (χ^2 $P = 0.75$) or poly-GA dipeptide repeat protein inclusions (χ^2 $P = 0.46$). Our findings suggest two non-exclusive pathogenic mechanisms: (i) functional depletion of RNA-processing proteins resulting in disruption of messenger RNA splicing; and (ii) licensing of expanded *C9orf72* pre-messenger RNA for nuclear export by inappropriate association with messenger RNA export adaptor protein(s) leading to cytoplasmic repeat associated non-ATG translation and formation of potentially toxic dipeptide repeat protein.

Received December 3, 2013. Revised March 3, 2014. Accepted April 2, 2014. Advance Access publication May 27, 2014

Keywords: amyotrophic lateral sclerosis; pathology; genetics; fluorescence imaging

Abbreviations: ALS = amyotrophic lateral sclerosis; FISH = fluorescence *in situ* hybridization

Introduction

Expanded GGGGCC repeats in intron 1 of the *C9orf72* gene represent the most common cause of familial amyotrophic lateral sclerosis (ALS) and familial frontotemporal degeneration (DeJesus-Hernandez *et al*., 2011; Renton *et al*., 2011), though how this genetic change results in neuronal injury is not yet understood. Three potential mechanisms have been proposed: (i) haploinsufficiency through disrupted expression of the expanded allele (DeJesus-Hernandez *et al*., 2011); (ii) RNA mediated gain-of-function toxicity by the transcribed expanded intronic sequence; and (iii) protein mediated gain-of-function toxicity by dipeptide repeat protein aberrantly translated from the repeat sequence by repeat associated non-ATG translation (Ash *et al*., 2013; Mori *et al*., 2013*b*). Evidence for haploinsufficiency is mixed; several groups have reported reduced expression of the *C9orf72* messenger RNA, but this finding is not consistent (Sareen *et al*., 2013). Furthermore no additional loss of function mutations have been found in the *C9orf72* gene (Harms *et al*., 2013) and we and others have shown that smaller repeat lengths, which are considered pathogenic (Byrne *et al*., 2013; Gomez-Tortosa *et al*., 2013), do not reduce transcription (Cooper-Knock *et al*., 2013; Xi *et al*., 2013). More evidence is being gathered for a gain-of-function toxicity mediated either by RNA foci formed from the expanded intron or through repeat associated non-ATG translation.

Recently, a number of studies reported that molecular phenotypes correlated with the presence of RNA foci (Donnelly *et al*., 2013; Lagier-Tourenne *et al*., 2013; Lee *et al*., 2013; Mizielinska *et al*., 2013; Sareen *et al*., 2013). Two of these studies corrected the observed phenotype by targeted degradation of the foci using antisense oligonucleotides (Donnelly *et al*., 2013; Sareen *et al*., 2013). One study suggested that foci burden in the frontal cortex positively correlated with disease severity in eight patients with *C9orf72* frontotemporal degeneration (Mizielinska *et al*., 2013). Two of these reports identified co-localization of RNA foci with various proteins (Donnelly *et al*., 2013; Sareen *et al*., 2013) and suggested that pathogenic sequestration might be occurring. A similar process has been observed in myotonic dystrophy type 1, another neuromuscular disease caused by an intronic expansion (Jiang *et al*., 2004). Previously two groups generated candidate binding partners of the GGGGCC repeat expansion, but did not include co-localization studies with RNA foci (Mori *et al*., 2013*a*; Xu *et al*., 2013). Further work to characterize protein binding partners of the RNA foci is required, particularly because many of the studies thus far are in disagreement as to the most important interactions.

Observations regarding toxicity of repeat associated non-ATG translation are still at an early stage: the produced dipeptide repeat protein appears to be toxic in a cell model (Zu *et al*., 2013), but levels of the aberrantly translated protein observed do not correlate with neurodegeneration in autopsy material (Mackenzie *et al*., 2013). An important question remains over the mechanism by which the transcribed repeat sequence is exported to the cytoplasm to allow repeat associated non-ATG translation. Clearly, normal control of messenger RNA nuclear export would be expected to inhibit this movement. However, several studies report cytoplasmic RNA foci in CNS tissue (Donnelly *et al*., 2013, Mizielinska *et al*., 2013).

We have used fluorescence *in situ* hybridization (FISH) to examine the abundance and location of RNA foci in cerebellum, where p62-positive protein inclusion pathology is characteristic of *C9orf72*+ disease (Cooper-Knock *et al*., 2012), and in motor neurons of the ventral horn. We also examined the relationship between RNA foci and characteristic neuropathology of *C9orf72*+ ALS: first, the loss of nuclear TDP-43 in motor neurons, which is the pathological hallmark of ALS (Neumann *et al*., 2006) and has been shown to correlate with neuronal loss (Brettschneider *et al*., 2013); and second, the presence of cytoplasmic aggregates containing dipeptide repeat protein, which are a hallmark of *C9orf72*+ disease (Ash *et al*., 2013; Mackenzie *et al*., 2013; Mori *et al*., 2013b). We have then identified protein binding partners of the RNA repeat expansion, initially in an *in vitro* RNA pulldown assay using both cerebellum and neuronal cell-line extracts, and then subsequently in CNS tissue from *C9orf72*+ patients with ALS by immunohistochemistry. Protein–RNA UV-crosslinking confirmed *in vitro* direct interactions with the repeat sequence. We add novel insights to this growing field and in particular, our focus on motor neurons from the ventral horn of the spinal cord has allowed us to characterize RNA foci and their interactions in the neuronal population most vulnerable to neurodegeneration in ALS.

It should be noted that other groups have observed RNA foci transcribed from the repeat sequence in an antisense direction consisting of a GGCCCC repeat (Gendron *et al*., 2013; Lagier-Tourenne *et al*., 2013; Mizielinska *et al*., 2013); antisense foci were not examined in this study.

Materials and methods

Human samples

The study was approved by the South Sheffield Research Ethics Committee and informed consent was obtained for all samples. Brain and spinal cord tissues were donated to the Sheffield Brain Tissue Bank for research with the consent of the next of kin. Immunohistochemistry and RNA FISH were performed on formalin fixed paraffin-embedded tissues from up to five *C9orf72*+ ALS cases, three *C9orf72*− ALS cases and three neurologically normal controls. Lymphoblastoid cells and fibroblasts from three *C9orf72*+ ALS cases, one *C9orf72*+ asymptomatic carrier, three *C9orf72*− ALS cases and three controls were used for RNA FISH. Lymphoblastoid cell lines were obtained from the Wellcome Trust/Motor Neurone Disease

Association ALS/MND UK DNA and Lymphoblastoid cell line Bank. Fibroblasts were obtained from the Sheffield MND Biosamples Bank.

RNA fluorescence *in situ* hybridization

A 5′ TYE-563-labelled LNA (16-mer fluorescent)-incorporated DNA probe was used against the sense RNA hexanucleotide repeat (Exiqon, Inc., batch number 607323). Slides with tissue, lymphoblastoid cells or fibroblasts were fixed in 4% paraformaldehyde for 10 min. Before use, formalin fixed paraffin-embedded tissue sections were deparaffinized. Slides were blocked with hybridization solution [50% formamide, 2 × saline sodium citrate (SSC), 100 mg/ml dextran sulphate, 50 mM sodium phosphate pH 7.0] for 3 h at 66°C and then incubated with 400 ng/ml of denatured probe in hybridization solution overnight at 66°C. After hybridization, slides were washed once in 2 × SSC/0.1% Tween-20 at room temperature and three times in 0.1 × SSC at 65°C. Slides were mounted with mounting medium containing DAPI (Vector Labs, Inc.). All solutions were made with DEPC-treated water.

Visualization of RNA foci

Primary visualization of foci was performed using a Leica SP5 confocal microscope system with a ×63/1.4 oil immersion objective lens. The presence of foci was assessed within a high resolution (1433 μm^2 per image, 511 × 511 pixels) *z*-stack made up of images at 0.13-μm intervals through the entire nuclear volume of the cell under consideration.

Biotinylated RNA pulldown assays

Total extracts were prepared by homogenizing and lysing cells/tissue in RNA-pulldown (RPD) lysis buffer [25 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT, 10% (v/v) glycerol, 0.5% (v/v) Triton^{TM} X-100]. Lysates were cleared by centrifugation and supernatants taken for experiments. Nuclear extracts from SH-SY5Y cells were prepared using the Dignam method (Dignam *et al*., 1983). We chose to use two methods of lysis because cell lysis has been shown to influence the composition of ribonucleoprotein complexes (Mili and Steitz, 2004).

AAAAUU$_5$ and GGGGCC$_5$ RNA molecules with 3′ biotin modifications were used to identify protein binding partners in pulldown assays. 60 μl aliquots of streptavidin sepharose (GE Healthcare) were blocked overnight on a spinning wheel at 4°C with RPD lysis buffer containing 2% bovine serum albumin. Total extracts were lysed in RPD lysis buffer whereas cerebellum homogenates and SH-SY5Y whole cell or nuclear extracts were mixed 1:1 with RPD lysis buffer (2 ×) supplemented with protease and RNase inhibitors. 1-2 mg of the appropriate total cellular or nuclear lysate was mixed with 15 μg biotin-labelled RNA, incubated at room temperature for 30 min and then on ice for 30 min. Mixtures were then transferred to a 6-cm petri dish and UV irradiated on ice at 0.3 J/cm^2 in a UV crosslinker (*Fisher*). Mixtures were then applied to blocked streptavidin sepharose and incubated at 4°C for 2 h with agitation. Following binding, beads were washed three times with RPD lysis buffer and then twice with RPD wash buffer (25 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT). Complexes were eluted by addition of RPD elution buffer (25 mM Tris pH 7.4, 25 mM NaCl, 1 mM EDTA) and 10 μg RNase A followed by agitation at room temperature for 30 min. Eluates were analysed by SDS-PAGE and proteins identified by mass spectrometry or western immunoblotting.

Mass spectrometry

In solution tryptic digestions were performed on the eluted fractions by the addition of 100 mM final concentration ammonium bicarbonate and 0.1% ProteaseMAX^{TM} surfactant. Trypsin was added to a mass ratio of (1:50) and incubated at 37°C overnight. Digestions were stopped with the addition of 1–2 μl glacial acetic acid and subsequently dried under vacuum. Tryptic digests were resuspended in 0.1% final concentration of trifluoroacetic acid. Five microlitres was used for liquid chromatography–mass spectrometry/mass spectrometry (LC–MS/MS) analysis. Peptides were separated using an UltiMate^{TM} 3000 RSLC nano liquid chromatography system (Dionex), using a 150 mm × 75 μm I.D. PepMap^{TM} reversed phase column (Dionex). Linear gradient elution was performed from 95% buffer A (0.1% formic acid) to 50 % buffer B (0.1% formic acid, 95 % acetonitrile) at a flow rate of 300 nl/min in 60 mins. MS/MS analysis was performed using a maXis UHR TOF mass spectrometer (Bruker Daltonics) using an automated acquisition approach. MS and MS/MS scans (m/z 50–2000) were acquired in positive ion mode. Lock mass calibration was performed using HP 1221.990364. Line spectra data were then processed into peak list by data analysis using the following settings. The sum peak finder algorithm was used for peak detection using a signal to noise ratio of 10, a relative to base peak intensity of 0.1% and an absolute intensity threshold of 100. Spectra were deconvoluted and the peak lists exported as Mascot Generic Files (MGF) and searched using Mascot 2.2 server (Matrix Science). The Swiss-Prot database (Swiss-Prot Release 10.5m5, 20 April 2010, 516604 sequences) was searched using the following parameters (analysis peptide tolerance = ±0.01 Da, MS/MS tolerance = ±0.01 Da and peptide charge 2+ and 3+). Search parameters were as follows: enzyme; trypsin; fixed modifications: carbamidomethyl (C); variable modifications: deamidation (NQ), oxidation (M); maximum missed cleavages: 1. Deamidation (NQ) were chosen as variable modifications. Additionally, we also used a peptide MOWSE score of <25 as a cut-off as calculated by Mascot. The false discovery rate was estimated to be 1% for peptide IDs after searching reverse databases. Protein identifications were based on a minimum of two unique peptides.

RNA-binding ultraviolet crosslinking assays

RNA-binding assays were carried out as described previously (Hautbergue *et al*., 2008, 2009). GGGGCC$_5$ RNA was 5′ end labelled with γ^{32}P-ATP using T4 polynucleotide kinase (Fermentas). Reaction mixes were made up in RNA binding buffer [15 mM HEPES pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 10% (v/v) glycerol, 0.05% (v/v) Tween-20] with 50 ng radiolabelled RNA and 5 μg purified recombinant protein. Mixes were incubated for 20 min at room temperature and 20 min on ice before being UV-irradiated on ice at full power. Complexes were analysed by SDS-PAGE and stained with Coomassie blue before being vacuum-dried and exposed on a phosphoimage screen.

Immunohistochemistry

The following antibodies were used for immunohistochemistry: anti-TDP-43 (Proteintech 10782-2-AP) anti-FUS (Novus NB100-2599), anti-hnRNP H1/F (Abcam ab10689), anti-hnRNP A1 (Abcam ab5832, 9H10 clone), anti-hnRNP D (Proteintech 12770-1-AP), anti-SRSF1 (phosphor, Abcam ab11826), anti-SRSF2 (Abcam ab30817),

anti-ALYREF (Sigma, clone 11G5) and anti-hnRNP C1/C2 (Abcam ab10294). Poly-(Gly-Ala) dipeptide repeat protein was detected using anti-GA antibodies (mouse, clone 5F2) as previously described (Mackenzie *et al*., 2013). Antigen retrieval was performed by 10–30-min microwave in EDTA at pH 8.0 for all antibodies except anti-SRSF1, anti-ALYREF and anti-TDP-43 where antigen retrieval involved microwave 10–20 min in trisodium citrate at pH 6.5, and for anti-hnRNP H/F where no specific antigen retrieval was performed. After incubation with the primary antibodies, slides were washed in PBS and incubated in species specific Alexa Fluor® 488-conjugated secondary antibodies.

Results

RNA fluorescence *in situ* hybridization

The presence of RNA foci clearly distinguished fibroblasts, lymphoblastoid cells and CNS tissue from *C9orf72*+ patients with ALS compared to *C9orf72*− patients with ALS and neurologically normal control subjects (Fig. 1A–D). To validate our RNA FISH methodology, discrete nuclear foci-like staining was quantified in a blinded study of 50 cerebellar granule neurons from each of nine cases: three *C9orf72*+ patients with ALS, three *C9orf72*− patients with ALS and three control subjects. In *C9orf72*+ tissue the average proportion of neurons containing nuclear RNA foci was 39% (range 21–63%); in three *C9orf72*− cases with ALS the average proportion of neurons containing foci-like staining was 1.6% (range 1.1–2.5%); in normal controls the average proportion of neurons containing foci-like staining was 1.4% (range 1.3–1.6%). Only seven foci-like objects were observed in 300 neurons from the six non-*C9orf72*+ cases and never was more than one focus-like object was observed in a single cell; in contrast the average rate in *C9orf72*+ tissue was two foci per cell. RNase treatment in fibroblasts ablated foci, illustrating that the labelled product is RNA and in agreement with previous studies (Fig. 1A).

It is noteworthy that RNA foci were identified in fibroblasts derived from an asymptomatic *C9orf72*+ carrier (Fig. 1C). In four *C9orf72*+ cases the proportion of foci+ cerebellar granule neurons was quantified and compared to the proportion of foci+ motor neurons in the ventral horn (Fig. 1E). More than 35 cells of each neuron-type were examined in each case. Three of the cases presented initially with ALS (Supplementary Fig. 1); in these patients the average proportion of foci+ neurons was significantly higher in the ventral horn (61% versus 27%, *t*-test $P < 0.05$). In the fourth case, who presented with frontotemporal degeneration and later developed ALS, the pattern was reversed (40% versus 63%). Foci were primarily nuclear, however, some cytoplasmic foci were also observed in fibroblasts, cerebellar granule cells and in motor neurons (Fig. 1C).

Identification of binding partners of the *C9orf72* repeat expansion

We generated 3′ biotinylated RNAs with the following sequences: 5′-[AAAAUU]$_5$-Bio-3′ and 5′-[GGGGCC]$_5$-Bio-3′. It has been demonstrated that the GGGGCC repeat expansion can form RNA G-quadruplexes *in vitro*, with the smallest repeating unit consisting of four repeats (Fratta *et al*., 2012; Reddy *et al*., 2013). To identify proteins interacting with the biotinylated RNAs, RNAs were preincubated with protein extracts and resulting complexes fixed by UV-irradiation. The RNA bait and bound proteins were captured using streptavidin sepharose and eluted after RNase A digestion. We used whole cell lysates of the human neuronal cell line SH-SY5Y, SH-SY5Y nuclear extract and dissected human cerebellum whole extract (Fig. 2A–C). Controls without RNA bait were processed in parallel (Fig. 2D). Eluted proteins were identified by mass spectrometry. In total, 103 unique proteins were identified that bind GGGGCC$_5$, the majority of which did not bind to AAAAUU$_5$ (Fig. 2E and Supplementary material).

Gene ontology (GO) enrichment analysis of each GGGGCC$_5$-derived list of bound proteins was carried out using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang da *et al*., 2009*a*, *b*). This yielded functional categories associated with aspects of messenger RNA metabolism including splicing and stabilization, and an RNA recognition motif-containing class (Supplementary Fig. 2). This was particularly striking in the list of GGGGCC RNA-binders isolated from nuclear extracts of the SH-SY5Y human neuronal cell line (Fig. 2F). Another strongly represented group was messenger RNA export adaptors, which promote nuclear export via remodelling of the NXF1/TAP export receptor (Hautbergue *et al*., 2008), including ALYREF and the shuttling splicing factors SRSF1 (SF2/ASF), SFRS3 (SRp20) and SFRS7 (9G8) (Walsh *et al*., 2010).

Cellular distribution of RNA foci and RNA recognition motif-containing proteins

We used confocal microscopy to validate *in vivo* some of the hits identified by mass spectrometry. For this purpose, eight well-described RNA recognition motif-containing proteins including splicing factors and messenger RNA nuclear export adaptors were prioritized and selected depending on available and efficacious commercial antibodies. The distribution of each protein relative to RNA foci was examined in approximately 200 cerebellar granule neurons and 50 motor neurons from a minimum of three *C9orf72*+ cases with ALS. Simultaneous co-staining was carried out in parallel in *C9orf72*− cases with ALS and neurologically normal control subjects. For all tested candidates, overall cellular protein distribution was not grossly different between *C9orf72*+ cases, *C9orf72*− cases and controls except for areas where co-localization was demonstrated. In cerebellar granule cells we demonstrated co-localization of hnRNP A1, hnRNP H1/F, ALYREF and SRSF2 with 27%, 30%, 26% and 33% of RNA foci, respectively (Fig. 3A–D). In motor neurons, the cell type most vulnerable to the neurodegenerative process in ALS, we demonstrated co-localization of hnRNP H1/F, ALYREF and SRSF2 with 19%, 29% and 30% RNA foci, respectively (Fig. 3E–G). In contrast, we were unable to detect any evidence of co-localization of other identified GGGGCC-binding partners SRSF1, FUS, hnRNP C or hnRNP D with sense foci in either the cerebellar granule layer or the ventral horn (Supplementary Fig. 3).

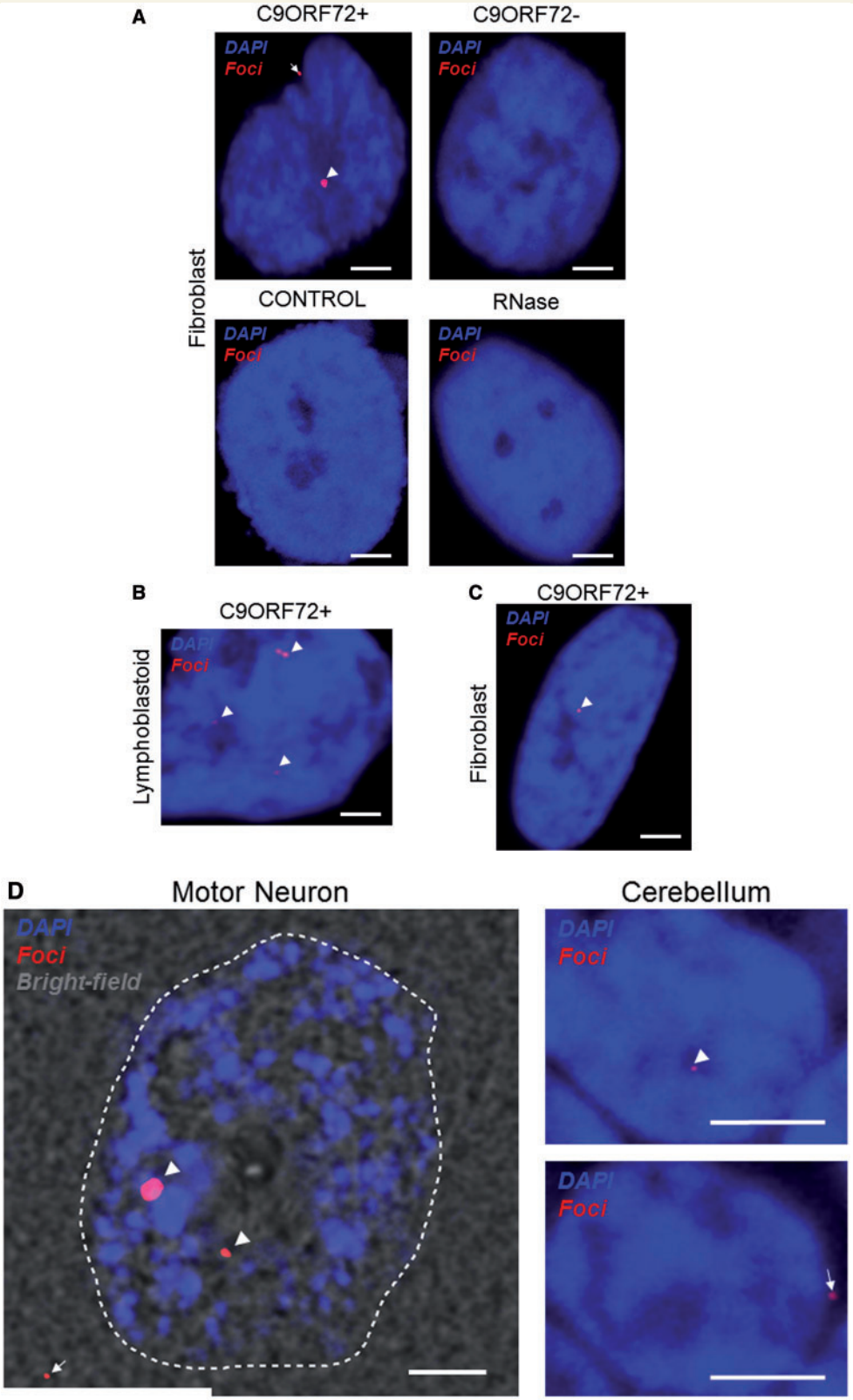


Figure 1 RNA FISH shows GGGGCC expanded RNA foci are found in peripheral cells and CNS tissue from *C9orf72* + patients but not from *C9orf72* − ALS cases or controls. RNA foci (arrowheads) are present in fibroblasts (**A**), lymphoblastoid cells (**B**) and CNS tissue (**D**) from *C9orf72* + patients with ALS, and in fibroblasts from a *C9orf72* + asymptomatic carrier (**C**). RNA foci are ablated by RNase treatment (**A**). RNA foci are predominantly nuclear but cytoplasmic foci are observed in peripheral cells and CNS tissue (**A** and **D**, arrows). Abundance of foci in cerebellar granule cells and motor neurons has been quantified (**E**), in those cases where the initial clinical presentation was ALS the proportion of foci + motor neurons is significantly higher ($*P < 0.05$). Scale bar = 3 μm. FTD = frontotemporal degeneration.

For six of the proteins identified in the mass spectrometry analysis, including those proteins observed to co-localize with RNA foci *in vivo*, specificity of interaction with the $(GGGGCC)_5$ RNA was assessed using RNA pull down assays from whole neuronal SH-SY5Y cell extract and western immunoblotting (Fig. 4A). Direct binding of some of these proteins to $(GGGGCC)_5$ RNA repeat was also confirmed in a UV-cross linking assay using radiolabelled RNA and recombinant proteins which were expressed and purified from *E. coli* (Fig. 4B).

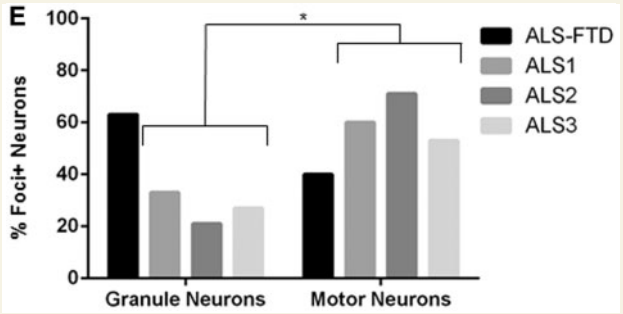
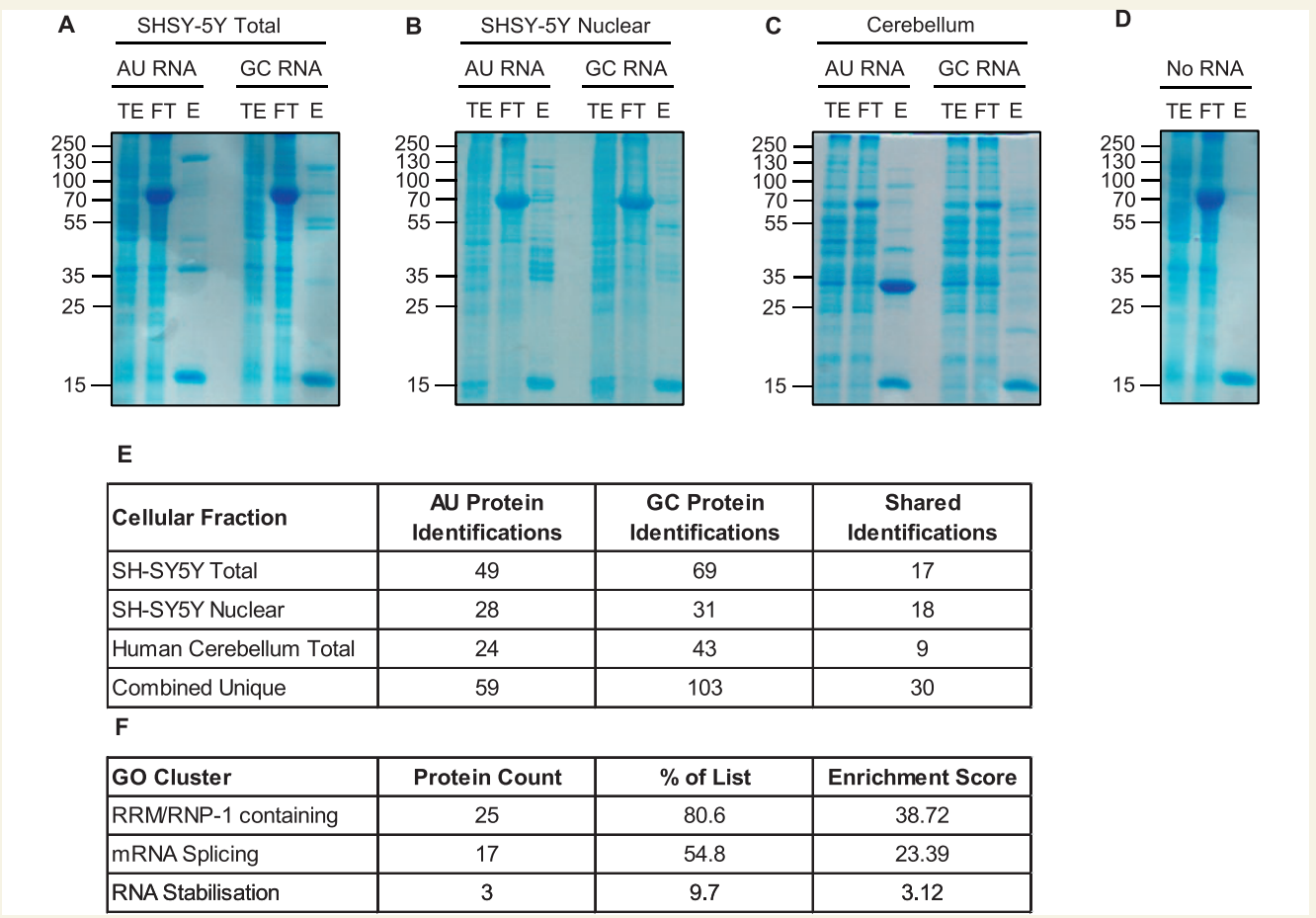


Figure 1 Continued.

We also examined the co-incidence of RNA foci with depletion of TDP-43 from the nuclei of motor neurons of *C9orf72* + patients with ALS. Mislocalization of TDP-43 is the pathophysiological hallmark of ALS (Neumann *et al*., 2006). All surviving motor neurons were examined in formalin fixed paraffin-embedded sections from three *C9orf72* + ALS cases. The majority of cells with nuclear depletion of TDP-43 contained nuclear RNA foci, but this was not significantly different to the proportion of cells with nuclear TDP-43 expression that contained RNA foci (66% versus 60%, χ^2 $P = 0.75$) (Supplementary Fig. 4).

In view of our prediction that the repeat sequence might sequester several proteins important for messenger RNA export, we wanted to explore the relationship between repeat associated non-ATG translated protein and RNA foci in specific neuronal populations. As expression of dipeptide repeat proteins is reported to be rare in the ventral horn of *C9orf72* + patients with ALS (Mackenzie *et al*., 2013), we chose to focus on cerebellar granule cells. Fifty per



E

Cellular Fraction	AU Protein Identifications	GC Protein Identifications	Shared Identifications
SH-SY5Y Total	49	69	17
SH-SY5Y Nuclear	28	31	18
Human Cerebellum Total	24	43	9
Combined Unique	59	103	30

F

GO Cluster	Protein Count	% of List	Enrichment Score
RRM/RNP-1 containing	25	80.6	38.72
mRNA Splicing	17	54.8	23.39
RNA Stabilisation	3	9.7	3.12

Figure 2 5′-$[AAAAUU]_5$- and 5′-$[GGGGCC]_5$ RNAs sequester distinct sets of proteins from human neuronal cell line fractions and dissected human cerebellar tissue. Pulldown assays using biotinylated RNAs (no RNA, 5′-$[AAAAUU]_5$- or 5′-$[GGGGCC]_5$) and extracts from total or nuclear fractions of SH-SY5Y cells, or human cerebellar tissue; I = input (1%); FT = flow through (1%); E = eluted (25%) (**A–D**). Mass spectrometry (MS) analysis of proteins co-purified with biotinylated RNAs (**E**). Gene ontology (GO) enrichment of SH-SY5Y nuclear hits (**F**).

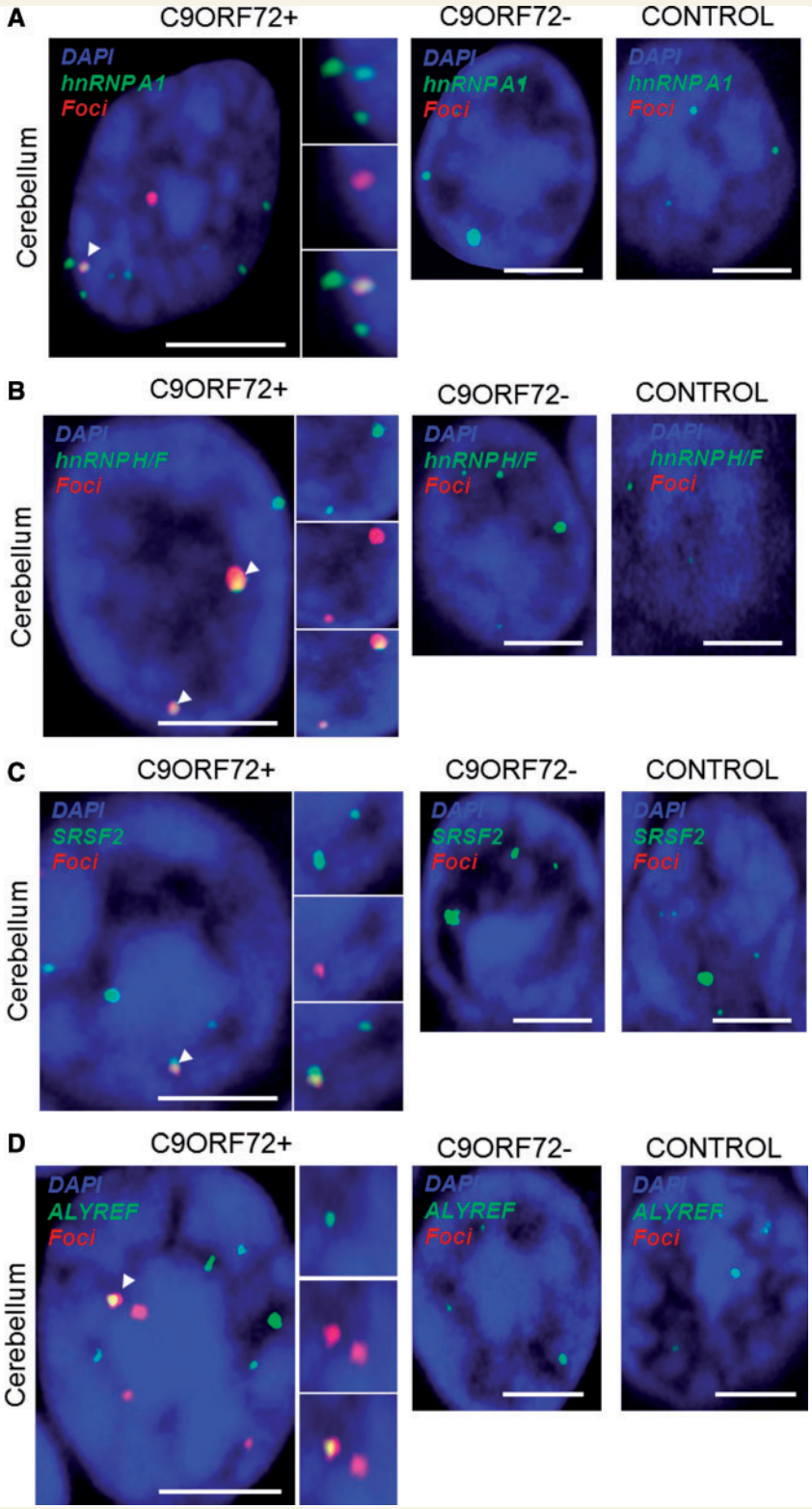


Figure 3 Combined RNA FISH and immunohistochemistry demonstrate co-localization of nuclear speckle components with RNA foci in CNS tissue. hnRNP A1 (**A**), hnRNP H1/F (**B**), SRSF2 (**C**) and ALYREF (**D**) are observed to co-localize with RNA foci (arrowheads) in cerebellar granule cells from *C9orf72* + patients with ALS. hnRNP H1/F (**E**), SRSF2 (**F**) and ALYREF (**G**) are observed to co-localize with RNA foci (arrowheads) in nuclei of motor neurons from *C9orf72* + patients with ALS. Co-localization events are enlarged and unmerged protein and RNA foci are shown for comparison. The normal staining pattern of the two proteins in *C9orf72* − cases with ALS and control subjects is included for comparison. Scale bar = 3 μm.

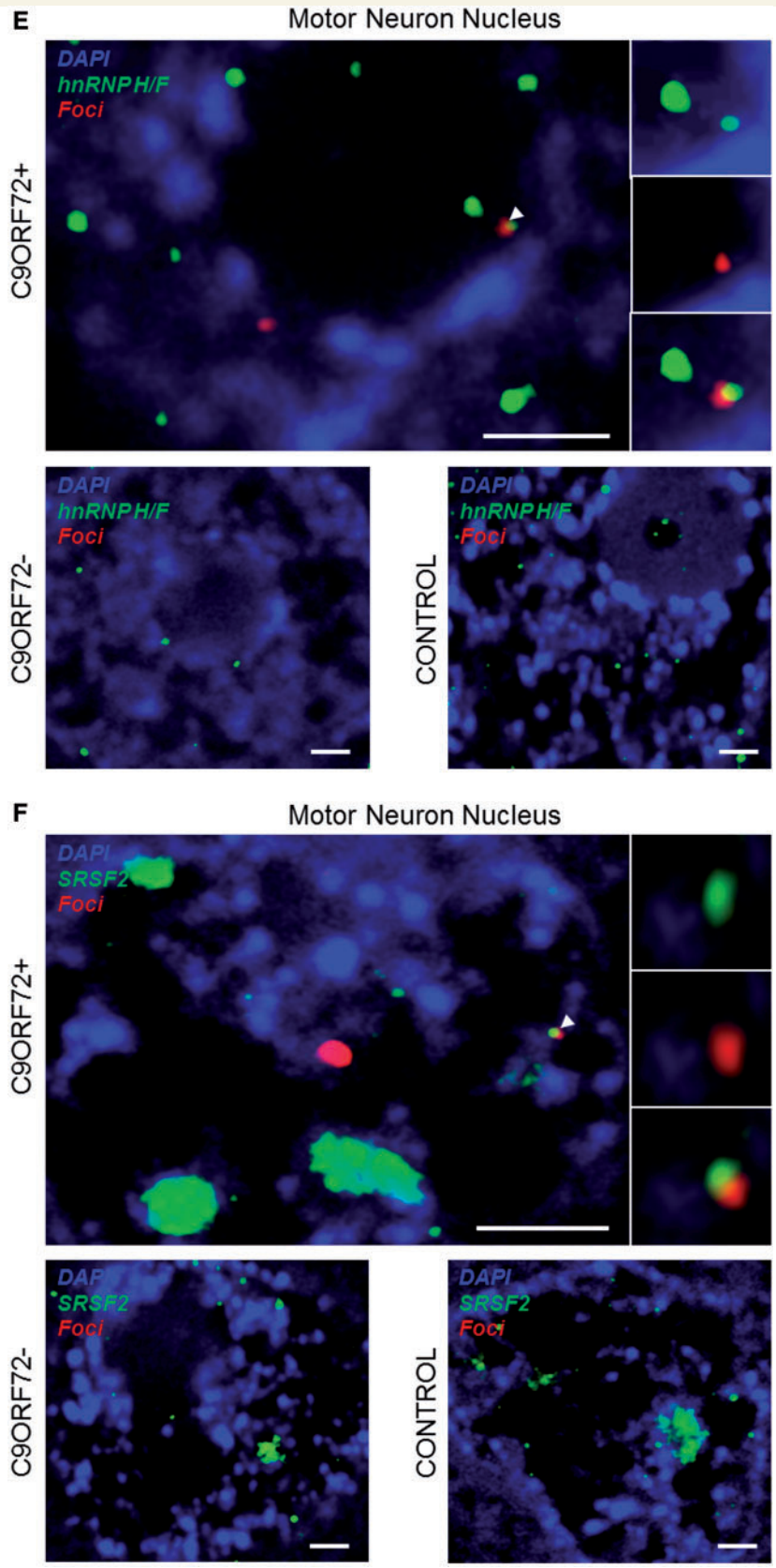


Figure 3 Continued.

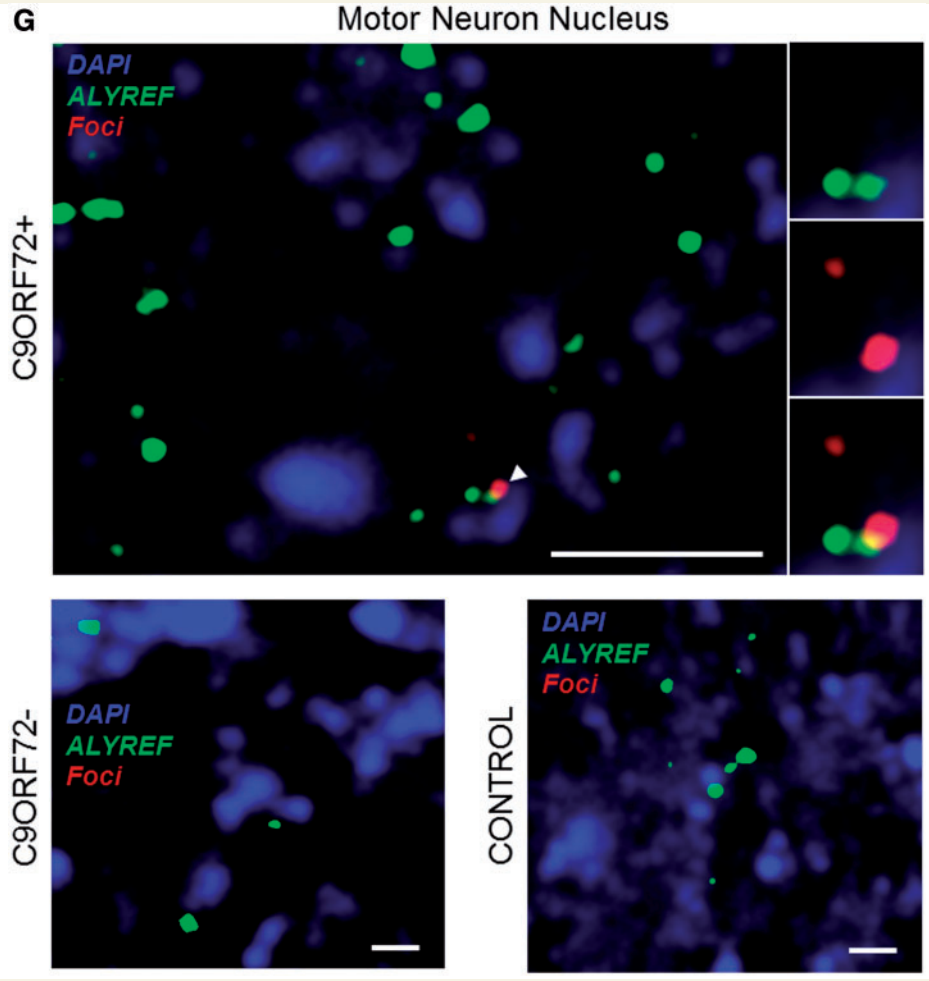


Figure 3 Continued.

cent of the neurons which stained for poly-GA, the most abundant dipeptide repeat protein, contained nuclear RNA foci; this was not significantly different to the proportion of neurons with nuclear RNA foci which did not stain for poly-GA (50% versus 40%, χ^2 $P = 0.46$) (Supplementary Fig. 5).

Discussion

There is an urgent need to understand the mechanisms of neuronal injury in *C9orf72* + disease. This genetic variant is the most common identified cause of ALS and frontotemporal degeneration. We and others (DeJesus-Hernandez *et al*., 2011; Donnelly *et al*., 2013; Lagier-Tourenne *et al*., 2013; Lee *et al*., 2013; Mizielinska *et al*., 2013; Sareen *et al*., 2013) have identified RNA foci formed from the intronic GGGGCC repeat sequence in peripheral cells and CNS tissue from *C9orf72* + patients. We have particularly focused on characterizing RNA foci within spinal motor neurons, which are the primary target of pathology in ALS. Indeed we have shown that RNA foci are present in a higher proportion of motor neurons of the ventral horn compared to cerebellar granule cells in patients where the initial clinical presentation was ALS; in a single patient where the initial clinical presentation was with extra-motor disease the opposite was true. This is consistent with toxicity initiated by RNA foci. However, this finding will require validation in a larger number of cases.

We have identified a number of putative binding partners of the RNA repeat expansion which are consistent with previous observations (Lee *et al*., 2013; Mori *et al*., 2013*a*; Sareen *et al*., 2013; Xu *et al*., 2013). Of the RNA recognition motif-containing proteins we found to be co-localized with RNA foci in *C9orf72* + tissue, hnRNP A1 (Sareen *et al*., 2013), hnRNP H1/F and SRSF2 (Lee *et al*., 2013) have been similarly observed by others. Interestingly, our study provides the first evidence for co-localization of RNA foci with the general messenger RNA nuclear export adaptor ALYREF (Stutz *et al*., 2000). Observed co-localization with RNA recognition motif-containing proteins was present in a relatively low percentage of RNA foci. We suggest that this is consistent with a process of dynamic sequestration. Indeed, irreversible binding of these candidates, many of which are key regulators of essential processes such as pre-messenger RNA splicing, is unlikely to be consistent with the relatively late age of disease onset seen in *C9orf72* + patients. The key pathogenic step may be downstream from protein sequestration by the

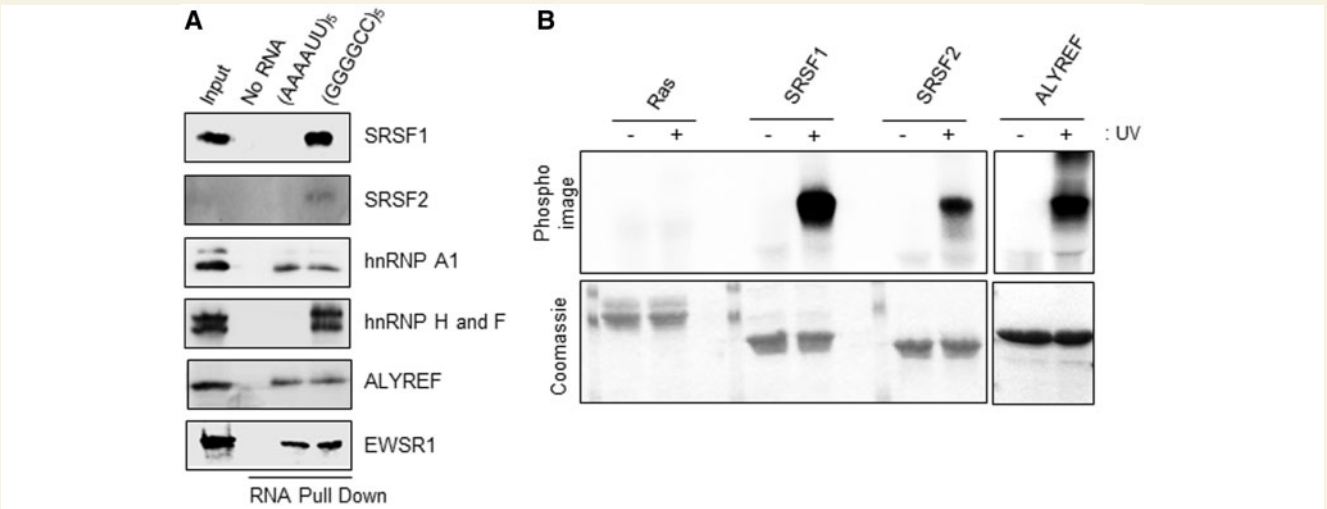


Figure 4 Identified RNA-binding candidates interact specifically and directly with $GGGGCC_5$ RNA. (**A**) Neuronal SH-SY5Y whole cell extract was incubated with either no RNA, AU-rich or GC-rich biotinylated RNA coated onto streptavidin beads before UV-cross linking. Bound proteins were eluted using RNase A and further identified using SDS-PAGE and western immunoblotting with the indicated antibodies. It is noted that the weak signal for SRSF2 is due to difficulty finding an antibody that is efficacious in western immunoblotting. The anti-hnRNP H1/F antibody recognizes both proteins, which are similar (Garneau *et al*., 2005). (**B**) Hexa-histidine-tagged recombinant SRSF1 11-196, GB1-tagged SRSF2 9-101 and ALYREF full length were expressed in *E. coli* and purified using metal ion affinity chromatography in 1 M NaCl containing buffers to remove potentially bound RNA from *E. coli* (*bottom*). $GGGGCC_5$ RNA was separately end-labelled with poly nucleotide kinase using [γ-^{32}P]-ATP, before incubation with purified proteins. RNA was covalently bound (+) or not (−) after UV irradiation. Absence of radioactive signal (*top*; PhosphoImage) in absence of UV irradiation demonstrates specificity of direct binding observed after UV treatment. All gels shown in the different panels were exposed simultaneously for the same amount of time (5 h).

expansion, such as export of the repeat expansion to enable repeat associated non-ATG translation or an accumulation of aberrant splicing events. Importantly we have confirmed a direct interaction *in vitro* between our protein candidates and the GGGGCC repeat RNA by UV-crosslinking.

SRSF2 is a well-known marker for nuclear speckles, nuclear domains implicated in the storage and supply of splicing factors to active transcription sites (Spector and Lamond , 2011). All of the proteins we have shown to co-localize with RNA foci, many of the binding partners identified in our RNA pulldown, and a number of the proteins implicated in genetic variants of ALS including TARDBP, EWSR1, FUS, HNRNPA1 and HNRNPA2B1, have been localized to nuclear speckles (Zhou *et al*., 2000; Saitoh *et al*., 2004; Casafont *et al*., 2009). Other neuromuscular diseases have been associated with depletion of normal components of nuclear speckles including myotonic dystrophy type 1 (Smith *et al*., 2007; Bengoechea *et al*., 2012). It is possible that disruption of the normal function of nuclear speckles, either by a direct mutation of one of the key protein components, or via RNA foci-mediated dynamic depletion of essential protein constituents, is a key pathogenic mechanism in ALS. Analysis of the transcriptome of pathologically affected neurons will be key to elucidating whether the interactions we have identified have a toxic effect through disruption of messenger RNA splicing.

We provide evidence for cytoplasmic RNA foci, not only in peripheral cells and in the cerebellar granule layer, but also in motor neurons from the ventral horn of the spinal cord. Cytoplasmic localization of RNA foci formed from an intronic repeat sequence in peripheral cells might be consistent with extrusion during mitosis. However, in non-dividing neurons of the cerebellum and ventral horn this is not a possibility. The alternative scenario relates to nuclear export of the transcribed GGGGCC repeat expansion. Our RNA pulldown screen for binding of the repeat expansion identified multiple messenger RNA export adapters including ALYREF (Stutz *et al*., 2000), SRSF1, SFRS3 and SFRS7 (Huang *et al*., 2003; Hargous *et al*., 2006; Tintaru *et al*., 2007). In the case of ALYREF we have also demonstrated co-localization with RNA foci by immunohistochemistry, and a direct interaction with the expansion by protein-RNA UV-crosslinking. An interesting possibility is that local enrichment of messenger RNA export adaptors onto *C9orf72* GGGGCC repeat pre-messenger RNA molecules overrides the normal nuclear retention of pre-messenger RNA, for example through an inappropriate interaction of ALYREF with the TAP/NXF1 nuclear export receptor. It seems unlikely that the RNA foci are exported intact, particularly because of their size and activity of DEAD box RNA helicases such as Dbp5/DDX19, on the cytoplasmic side of the nuclear pore which would be expected to unwind G-quadruplex structures (Linder, 2008). However, it is conceivable that aberrantly expanded *C9orf72* pre-messenger RNA molecules are exported from the nucleus and reform into foci within the cytoplasm.

Nuclear export of GGGGCC repeat RNA is likely to be a key step leading to repeat associated non-ATG translation in the cytoplasm. If dipeptide repeat proteins formed in this manner are

eventually identified as the key mediator of pathogenicity in *C9orf72* + disease then blocking this export represents an attractive therapeutic target. One report has suggested that the production of repeat associated non-ATG translated protein is mutually exclusive to the presence of RNA foci (Donnelly *et al*., 2013). In contrast, we found an equal proportion of poly-GA staining in neurons that did or did not contain RNA foci.

We did not observe a significant correlation between nuclear loss of TDP-43 and the presence of RNA foci. This does not mean that RNA foci are not instrumental in the disease pathogenesis, but may reflect the fact that they occur significantly upstream of TDP-43 mislocalization. In this regard it is important to note that we and others (Lagier-Tourenne *et al*., 2013) have identified RNA foci in fibroblasts derived from asymptomatic *C9orf72* + carriers.

We await confirmation of our findings by other groups. We have suggested two ways in which the interactions identified may be pathogenic: (i) through disruption of the normal function of factors involved in nuclear speckles and thus messenger RNA splicing; and (ii) through inappropriate licensing of the transcribed *C9orf72* expansion for nuclear export thereby facilitating repeat associated non-ATG translation. Either or both may be important, but it should be noted that if inappropriate licensing of RNA foci for export is a key pathogenic step, then overexpression of the sequestered protein will not be of therapeutic benefit and may even have an adverse effect. On the contrary if loss of the normal function of these proteins is most important, then increasing the nuclear expression of proteins sequestered by the expansion may be of value as a neuroprotective strategy.

Acknowledgements

We are grateful to all of the patients with ALS and their family members who donated biosamples for research.

Funding

We acknowledge grants from the EU Framework 7 (Euromotor No259867) and the SOPHIA project (funded by JPND and MRC) to P.J.S. and J.K. P.J.S. is an NIHR Senior Investigator. J.C.K. and J.R.H. are supported by MND Association/MRC Lady Edith Wolfson Fellowship awards ([MR/K003771/1] and [G0 800380] respectively). Biosample collection was supported by the MND Association and the Wellcome Trust (P.J.S.). M.J.D. acknowledges support from the Engineering and Physical Sciences Research Council (UK) and the Biotechnology and Biological Sciences Research Council (UK). D.E. was supported by the Helmholtz Young Investigator program HZ-NG-607.

Supplementary material

Supplementary material is available at *Brain* online.

References

Ash PE, Bieniek KF, Gendron TF, Caulfield T, Lin WL, Dejesus-Hernandez M, et al. Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron 2013; 77: 639–46.

Bengoechea R, Tapia O, Casafont I, Berciano J, Lafarga M, Berciano MT. Nuclear speckles are involved in nuclear aggregation of PABPN1 and in the pathophysiology of oculopharyngeal muscular dystrophy. Neurobiol Dis 2012; 46: 118–29.

Brettschneider J, Del Tredici K, Toledo JB, Robinson JL, Irwin DJ, Grossman M, et al. Stages of pTDP-43 pathology in amyotrophic lateral sclerosis. Ann Neurol 2013; 74: 20–38.

Byrne S, Heverin M, Elamin M, Walsh C, Hardiman O. Intermediate repeat expansion length in C9orf72 may be pathological in amyotrophic lateral sclerosis. Amyotroph Lateral Scler Frontotemporal Degener 2013; 15: 148–50.

Casafont I, Bengoechea R, Tapia O, Berciano MT, Lafarga M. TDP-43 localizes in mRNA transcription and processing sites in mammalian neurons. J Struct Biol 2009; 167: 235–41.

Cooper-Knock J, Hewitt C, Highley JR, Brockington A, Milano A, Man S, et al. Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. Brain 2012; 135: 751–64.

Cooper-Knock J, Higginbottom A, Connor-Robson N, Bayatti N, Bury JJ, Kirby J, et al. C9ORF72 transcription in a frontotemporal dementia case with two expanded alleles. Neurology 2013; 81: 1719–21.

DeJesus-Hernandez M, Mackenzie I, Boeve B, Boxer A, Baker M, Rutherford N, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-Linked FTD and ALS. Neuron 2011; 72: 245–56.

Dignam JD, Lebovitz RM, Roeder RG. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res 1983; 11: 1475–89.

Donnelly CJ, Zhang PW, Pham JT, Heusler AR, Mistry NA, Vidensky S, et al. RNA Toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. Neuron 2013; 80: 415–28.

Fratta P, Mizielinska S, Nicoll AJ, Zloh M, Fisher EM, Parkinson G, et al. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. Sci Rep 2012; 2: 1016.

Garneau D, Revil T, Fisette JF, Chabot B. Heterogeneous nuclear ribonucleoprotein F/H proteins modulate the alternative splicing of the apoptotic mediator Bcl-x. J Biol Chem 2005; 280: 22641–50.

Gendron TF, Bieniek KF, Zhang YJ, Jansen-West K, Ash PE, Caulfield T, et al. Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. Acta Neuropathol 2013; 126: 829–44.

Gomez-Tortosa E, Gallego J, Guerrero-Lopez R, Marcos A, Gil-Neciga E, Sainz MJ, et al. C9ORF72 hexanucleotide expansions of 20–22 repeats are associated with frontotemporal deterioration. Neurology 2013; 80: 366–70.

Hargous Y, Hautbergue GM, Tintaru AM, Skrisovska L, Golovanov AP, Stevenin J, et al. Molecular basis of RNA recognition and TAP binding by the SR proteins SRp20 and 9G8. EMBO J 2006; 25: 5126–37.

Harms MB, Cady J, Zaidman C, Cooper P, Bali T, Allred P, et al. Lack of C9ORF72 coding mutations supports a gain of function for repeat expansions in amyotrophic lateral sclerosis. Neurobiol Aging 2013; 34: 2234.e2213–39.

Hautbergue GM, Hung ML, Golovanov AP, Lian LY, Wilson SA. Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. Proc Natl Acad Sci USA 2008; 105: 5154–9.

Hautbergue GM, Hung ML, Walsh MJ, Snijders AP, Chang CT, Jones R, et al. UIF, a New mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. Curr Biol 2009; 19: 1918–24.

Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009a; 37: 1–13.

Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009b; 4: 44–57.

Huang Y, Gattoni R, Stévenin J, Steitz JA. SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. Mol Cell 2003; 11: 837–43.

Jiang H, Mankodi A, Swanson MS, Moxley RT, Thornton CA. Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. Hum Mol Genet 2004; 13: 3079–88.

Lagier-Tourenne C, Baughn M, Rigo F, Sun S, Liu P, Li HR, et al. Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. Proc Natl Acad Sci USA 2013; 110: E4530–9.

Lee YB, Chen HJ, Peres JN, Gomez-Deza J, Attig J, Stalekar M, et al. Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. Cell Rep 2013; 5: 1178–86.

Linder P. mRNA export: RNP remodeling by DEAD-box proteins. Curr Biol 2008; 18: R297–9.

Mackenzie IR, Arzberger T, Kremmer E, Troost D, Lorenzl S, Mori K, et al. Dipeptide repeat protein pathology in C9ORF72 mutation cases: clinico-pathological correlations. Acta Neuropathol 2013; 126: 859–79.

Mili S, Steitz JA. Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. RNA 2004; 10: 1692–4.

Mizielinska S, Lashley T, Norona FE, Clayton EL, Ridler CE, Fratta P, et al. C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. Acta Neuropathol 2013; 126: 845–57.

Mori K, Lammich S, Mackenzie IR, Forne I, Zilow S, Kretzschmar H, et al. hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with C9orf72 mutations. Acta Neuropathol 2013a; 125: 413–23.

Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E, et al. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science 2013b; 339: 1335–8.

Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science 2006; 314: 130–3.

Reddy K, Zamiri B, Stanley SY, Macgregor RB Jr, Pearson CE. The disease-associated r(GGGGCC)n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. J Biol Chem 2013; 288: 9860–6.

Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 2011; 72: 257–68.

Saitoh N, Spahr CS, Patterson SD, Bubulya P, Neuwald AF, Spector DL. Proteomic analysis of interchromatin granule clusters. Mol Biol Cell 2004; 15: 3876–90.

Sareen D, O'Rourke JG, Meera P, Muhammad AK, Grant S, Simpkinson M, et al. Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. Sci Transl Med 2013; 5: 208ra149.

Smith KP, Byron M, Johnson C, Xing Y, Lawrence JB. Defining early steps in mRNA transport: mutant mRNA in myotonic dystrophy type I is blocked at entry into SC-35 domains. J Cell Biol 2007; 178: 951–64.

Spector DL, Lamond AI. Nuclear speckles. Cold Spring Harb Perspecti Biol 2011; 3: pii: a000646.

Stutz F, Bachi A, Doerks T, Braun IC, Seraphin B, Wilm M, et al. REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. RNA 2000; 6: 638–50.

Tintaru AM, Hautbergue GM, Hounslow AM, Hung ML, Lian LY, Craven CJ, et al. Structural and functional analysis of RNA and TAP binding to SF2/ASF. EMBO Rep 2007; 8: 756–62.

Walsh MJ, Hautbergue GM, Wilson SA. Structure and function of mRNA export adaptors. Biochem Soc Trans 2010; 38: 232–6.

Xi Z, Zinman L, Moreno D, Schymick J, Liang Y, Sato C, et al. Hypermethylation of the CpG island near the GC repeat in ALS with a C9orf72 expansion. Am J Hum Genet 2013; 92: 981–9.

Xu Z, Poidevin M, Li X, Li Y, Shu L, Nelson DL, et al. Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. Proc Natl Acad Sci USA 2013; 110: 7778–83.

Zhou Z, Luo MJ, Straesser K, Katahira J, Hurt E, Reed R. The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. Nature 2000; 407: 401–5.

Zu T, Liu Y, Banez-Coronel M, Reid T, Pletnikova O, Lewis J, et al. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. Proc Natl Acad Sci USA 2013; 110: E4968–77.

Acta Neuropathol (2015) 130:63–75
DOI 10.1007/s00401-015-1429-9

ORIGINAL PAPER

Antisense RNA foci in the motor neurons of *C9ORF72*-ALS patients are associated with TDP-43 proteinopathy

Johnathan Cooper-Knock[1] · Adrian Higginbottom[1] · Matthew J. Stopford[1] · J. Robin Highley[1] · Paul G. Ince[1] · Stephen B. Wharton[1] · Stuart Pickering-Brown[2] · Janine Kirby[1] · Guillaume M. Hautbergue[1] · Pamela J. Shaw[1]

Received: 19 February 2015 / Revised: 8 April 2015 / Accepted: 20 April 2015 / Published online: 6 May 2015

Abstract GGGGCC repeat expansions of *C9ORF72* represent the most common genetic variant of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia. We and others have proposed that RNA transcribed from the repeat sequence is toxic via sequestration of RNA-binding factors. Both GGGGCC-repeat (sense) and CCCCGG-repeat (antisense) molecules are detectable by fluorescence in situ hybridisation as RNA foci, but their relative expression pattern within the CNS and contribution to disease has not been determined. Blinded examination of CNS biosamples from ALS patients with a repeat expansion of *C9ORF72* showed that antisense foci are present at a significantly higher frequency in cerebellar Purkinje neurons and motor neurons, whereas sense foci are present at a significantly higher frequency in cerebellar granule neurons. Consistent with this, inclusions containing sense or antisense derived dipeptide repeat proteins were present at significantly higher frequency in cerebellar granule neurons or motor neurons, respectively. Immunohistochemistry and UV-crosslinking studies showed that sense and antisense RNA molecules share similar interactions with SRSF2, hnRNP K, hnRNP A1, ALYREF, and hnRNP H/F. Together these data suggest that, although sense and antisense RNA molecules might be expected to be equally toxic via their shared protein binding partners, distinct patterns of expression in various CNS neuronal populations could lead to relative differences in their contribution to the pathogenesis of neuronal injury. Moreover in motor neurons, which are the primary target of pathology in ALS, the presence of antisense foci (χ^2, $p < 0.00001$) but not sense foci (χ^2, $p = 0.75$) correlated with mislocalisation of TDP-43, which is the hallmark of ALS neurodegeneration. This has implications for translational approaches to *C9ORF72* disease, and furthermore interacting RNA-processing factors and transcriptional activators responsible for antisense versus sense transcription might represent novel therapeutic targets.

Keywords *C9ORF72* · Amyotrophic lateral sclerosis · RNA foci · Dipeptide repeat protein · Immunohistochemistry

Introduction

GGGGCC hexanucleotide repeat expansions in *C9ORF72* represent the most common genetic variant of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) [7, 27]. The mechanism of pathogenesis is unknown, but it has been suggested that a gain-of-function toxicity may be mediated via sequestration of RNA recognition motif (RRM) containing proteins by RNA foci [2, 3]. It has been observed that RNA foci are formed, not only from sense, but also from antisense transcription of the repeat expansion [7, 16, 21]. The relative contribution of GGGGCC-repeat (sense) and CCCCGG-repeat (antisense) RNA molecules to disease pathogenesis is unknown, but is likely to have significant implications for subsequent translational

Electronic supplementary material The online version of this article (doi:10.1007/s00401-015-1429-9) contains supplementary material, which is available to authorized users.

✉ Pamela J. Shaw
pamela.shaw@sheffield.ac.uk

[1] Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, 385A Glossop Road, Sheffield S10 2HQ, UK

[2] Institute of Brain, Behaviour and Mental Health, 2.014 AV Hill Building, University of Manchester, Manchester M13 9PT, UK

research. Work by Haeusler et al. [11] recently suggested that, with a small number of exceptions, the protein binding partners of the two species of RNA foci are similar.

Another suggested mechanism of pathogenesis is direct toxicity of one or more of five dipeptide repeat proteins (DPRs) translated in different reading frames from either the sense [23] or antisense [24] RNA molecules. Poly(Gly-Ala) (GA) and poly(Gly-Arg) (GR) are translated from sense RNA molecules; poly(Pro-Ala) (PA) and poly(Pro-Arg) (PR) are translated from the antisense RNA molecules and poly(Pro-Gly) (PG) is translated from both molecules. Several recent studies have described how these proteins might disrupt ribosomal RNA biogenesis and pre-mRNA splicing [15, 22] or form toxic aggregates [20]. If DPRs are key to pathogenesis, then aberrant nuclear export of repeat RNA sequences, which is necessary to facilitate access to cytoplasmic translation machinery, may be an attractive therapeutic target. We have previously identified interactions between sense RNA repeat sequences and mRNA export adaptor proteins which might have a role in inappropriate licencing for nuclear export [3].

We conducted extensive immunohistochemistry (IHC) in tissue from *C9ORF72*-ALS cases to determine the distribution of each species of RNA foci within various CNS neuronal populations known to degenerate in *C9ORF72*-disease [19]. Blinded examination of serial sections showed that antisense foci are present at a higher frequency in cerebellar Purkinje neurons and motor neurons, whereas sense foci are present at a higher frequency in cerebellar granule neurons. Similar examination in neuronal populations of the hippocampal dentate gyrus and CA4 subfield did not reveal a consistent distinction, with significant variability between cases. Moreover, neuronal inclusions containing DPRs translated from sense RNA are present at a higher frequency in cerebellar granule neurons, whereas neuronal inclusions containing DPRs translated from antisense RNA are present at a higher frequency in motor neurons. Notably, motor neurons are the primary target of pathology in ALS. Furthermore we examined the distribution of RRM-containing proteins predicted to bind one or both of sense and antisense foci with specific attention to colocalisation with antisense RNA foci. Direct and specific binding to the antisense/sense repeat sequence was examined by UV crosslinking using purified recombinant proteins. Finally, we studied the relative association of each species of RNA foci with the hallmark of ALS neurodegeneration, namely mislocalisation of TDP-43 in motor neurons [25]. We add novel insights to this field—in particular our focus on neuropathology has allowed us to contextualize the sense and antisense RNA foci within framework of the human disease.

Materials and methods

Human samples

This study was approved by the South Sheffield Research Ethics Committee and informed consent was obtained for all samples. Brain and spinal cord tissues were donated to the Sheffield Brain Tissue Bank for research, with the consent of the next of kin. IHC and RNA fluorescence in situ hybridisation (FISH) were performed on formalin fixed paraffin-embedded (FFPE) tissues from eight *C9ORF72*+ patients with ALS and/or FTD, three non-*C9ORF72* ALS patients, and three neurologically normal controls. Clinical features of cases examined are summarized in Table 1.

RNA FISH

A 5′ TYE-563-labelled LNA (16-mer fluorescent)-incorporated DNA probe was used against the sense (Exiqon, Inc.; batch number 607323) and the antisense RNA hexanucleotide repeat (Exiqon, Inc.; batch number 610331). Slides were prepared and RNA foci were visualised as described previously [3] using a Leica SP5 confocal microscope system with a ×63/1.4 oil immersion objective lens. Briefly prehybridisation was followed by overnight hybridization

Table 1 Clinical details of *C9ORF72*+ cases used in pathological analysis

Case	Phenotype	Sex (M/F)	Age at onset (Years)	Disease duration (Months)	Site of onset	Post-mortem delay (h)
1	ALS-FTD	F	63	43	Cognitive	24
2	ALS	F	56	43	Limb	32
3	ALS	M	69	38	Limb	~96
4	ALS	F	61	40	Bulbar	7
5	ALS	F	58	7	Limb	2
6	ALS	M	62	20	Bulbar	~48
7	ALS	F	50	28	Bulbar	22
8	FTD	F	58	36	Cognitive	N/A

Case numbers are matched in Tables 1, 2, and 3

at 66 °C in a humid atmosphere. A single wash at room temperature with 2 × SSC/0.1 % Tween-20 preceded three washes at 65 °C with 0.1 × SSC. Slides were then mounted in DAPI Vectashield or processed further for dual staining of RNA and protein.

RNA-binding UV-crosslinking assays

RNA-binding assays were carried out as described previously [12, 13]. Recombinant proteins were expressed and purified from *E.coli* (Supplementary Table 1). Magoh, SRSF2 9-101, ALYREF, hnRNP A1-like2, hnRNP K, and hnRNP F were expressed in *E. coli* and purified by Ion Metal Affinity Chromatography in 1 M NaCl containing buffers to remove potentially bound RNA from E. coli. hnRNP K was further purified by ion exchange chromatography using a Mono Q column (GE healthcare).

$(GGCCCC)_5$ and $(CCCCGG)_5$ RNAs were 5′ end labelled with [ɣ32P]-ATP using T4 polynucleotide kinase (Fermentas). Reaction mixes were made up in RNA binding buffer [15 mM HEPES pH 7.5, 500 mM NaCl, 5 mM MgCl2, 10 % (v/v) glycerol, 0.05 % (v/v) Tween-20] with 50 ng radiolabelled RNA and 2 µg purified recombinant protein. Mixes were incubated for 10 min at room temperature before being UV irradiated on ice at full power. Complexes were analysed by SDS/PAGE and stained with Coomassie blue before being vacuum dried and exposed on a phosphoimage screen.

Immunohistochemistry

The following antibodies were used for IHC anti-TDP-43 (Proteintech 10782-2-AP), anti-hnRNP H/F (Abcam ab10689), anti-hnRNP A1 (Abcam ab5832, 9H10 clone), anti-SRSF2 (Abcam ab11826), anti-ALYREF (Sigma, clone 11G5), anti-nucleolin (Proteintech 10556-1-AP), and anti-hnRNP K (Abcam ab52600). Poly-GA was detected with anti-GA antibodies (mouse, clone 5F2) as previously described [18]. Poly-GR, poly-PA, poly-PR, and poly-PG were detected with antibodies provided by Stuart Pickering-Brown (Proteintech, Manchester, UK). For anti-hnRNP A1 and anti-SRSF2, antigen retrieval was performed by microwaving for 10–30 min in EDTA at pH 8.0. For all other antibodies, antigen retrieval involved 10–20 min microwave in trisodium citrate at pH 6.5 except for anti-hnRNP H/F where no specific antigen retrieval was performed. After incubation with the primary antibodies overnight at 4 °C in DEPC-treated PBS/5 % BSA slides were washed in DEPC PBS and incubated in fluorescent species-specific secondary antibodies. When dual staining of protein and RNA was performed, RNA FISH was performed first after which slides were immediately transferred to PBS/5 % BSA for protein staining.

Results

Relative distribution of sense and antisense RNA foci

The frequency of sense and antisense RNA foci was determined in five neuronal populations: Purkinje and granule neurons in the cerebellum, motor neurons of the spinal cord ventral horn, and neurons of the hippocampal dentate gyrus and CA4 subfield. These neuronal populations were chosen as they all exhibit neurodegeneration in *C9ORF72*-ALS and are characteristic of both motor (motor neurons) and extra-motor (cerebellum and hippocampus) pathology [4]. Sequential sections of tissue from *C9ORF72*-ALS cases, non-*C9ORF72* ALS cases, and controls were examined for RNA foci in a blinded manner. No RNA foci were observed in tissue from controls and non-*C9ORF72* ALS cases. Forty Purkinje neurons, forty motor neurons, >200 granule neurons, >150 dentate gyrus neurons, and >100 CA4 subfield neurons were evaluated from four C9ORF72+ ALS and/or FTD cases. The average frequency of sense and antisense foci per cell is shown in Table 2 (raw data are shown in Supplementary Table 2). Comparison between cases showed that the frequency of sense and antisense RNA foci was positively correlated in all neuronal populations i.e., cases with more sense foci per cell also had more antisense foci per cell. The exception to this was the dentate gyrus neurons where case-to-case variability was smallest (Pearson correlation coefficient: cerebellar Purkinje neurons 0.99, cerebellar granule neurons 0.6, motor neurons 0.2, CA4 subfield neurons 0.65) (Table 2). In the cerebellar populations and motor neurons but not hippocampal neurons, there was a difference between the frequency of antisense and sense foci which was consistent between cases (representative images are shown in Fig. 1a). To determine whether this difference was statistically significant, the foci count was modelled as a Poisson distribution and performing a likelihood-ratio test revealed that, within each individual case, the frequency of antisense compared to sense RNA foci was significantly higher in Purkinje neurons (likelihood-ratio test $p < 0.05$) and motor neurons (likelihood-ratio test $p < 0.05$), but significantly lower in cerebellar granule neurons (likelihood-ratio test $p < 0.05$). The fact that sense and antisense foci were relatively more abundant in different neuronal populations is against an artefact caused by differences in affinity of RNA FISH probes.

As reported for sense RNA foci [3], we observed cytoplasmic antisense RNA foci even in post-mitotic mature cells such as motor neurons (Fig. 1b).

Relative distribution of DPRs derived from sense and antisense RNA sequences

Staining of poly-GA, poly-GR, poly-PA, poly-PR, and poly-PG protein was studied in cerebellar granule neurons and motor neurons from three *C9ORF72*-ALS cases. More than 1000 granule neurons and approximately 50 motor

Table 2 Mean and standard deviation (SD) of number of sense and antisense RNA foci per nucleus in Purkinje neurons, granule neurons, motor neurons, dentate gyrus neurons, and CA4 subfield neurons in four *C9ORF72*-ALS patients

Case	Antisense (mean)	Antisense (SD)	Sense (mean)	Sense (SD)	*p* value
Purkinje neurons					
1	26.40	20.3	6.40	17.4	2.37E−14
2	4.30	4.19	1.10	2.18	0.002
3	4.60	4.50	1.30	1.95	0.002
4	6.30	5.54	1.40	1.51	6.88E−05
Granule neurons					
1	0.00	0.00	0.56	1.20	1.37E−12
2	0.03	0.17	1.10	1.41	9.63E−18
3	0.01	0.10	0.34	0.93	1.17E−07
4	0.02	0.14	0.40	0.82	3.1E−07
Motor neurons					
1	14.90	24.5	1.50	1.65	5.65E−14
2	3.00	4.06	1.00	1.25	0.02
3	3.33	3.78	1.00	0.76	0.02
4	5.40	6.52	2.44	3.40	0.02

Case	Antisense (mean)	Antisense (SD)	Sense (mean)	Sense (SD)
Dentate gyrus neurons				
1	0.65	1.90	0.88	2.34
6	1.65	3.35	0.95	1.81
7	0.89	2.42	1.39	2.23
8	1.63	5.34	0.91	1.76
CA4 subfield neurons				
1	10.3	14.7	6.55	8.68
6	3.17	6.51	1.33	1.67
7	0.50	0.97	2.72	4.90
8	6.31	10.8	9.82	13.8

In each case, antisense RNA foci are significantly more numerous in Purkinje neurons and motor neurons (likelihood-ratio test $p < 0.05$) but significantly less numerous in granule neurons (likelihood-ratio test $p < 0.05$)

neurons were examined in a blinded experiment. Neuronal inclusions containing sense RNA derived DPRs were only observed in granule neurons, whereas inclusions containing antisense RNA derived DPRs were only observed in motor neurons (Fig. 2a). Above background staining for poly-PG was not observed in any cells. In motor neurons, inclusions were predominantly nuclear, but in granule neurons inclusions were predominantly cytoplasmic (Fig. 2a). Background staining was examined in control and non-*C9ORF72*-ALS cases.

As a further validation, and to extend, the conclusions of this study to a larger number of cases, dual staining of poly-GA, and poly-PA protein were examined in a further blinded experiment. Approximately 1000 granule neurons and 50 motor neurons were studied from six C9ORF72+ ALS and/or FTD cases including three cases not utilized in the earlier analysis. In each case, the correct protein was determined based on the frequency of observed inclusions (Fig. 2b). Modelling the number of neuronal inclusions as a Poisson distribution and performing a likelihood-ratio test revealed that the frequency of poly-GA inclusions was significantly higher in granule neurons (likelihood-ratio test $p < 0.01$) and the frequency of poly-PA inclusions was significantly higher in motor neurons (likelihood-ratio test $p < 0.01$). The average frequencies of inclusions containing poly-GA and poly-PA protein are shown in Table 3 (raw data are shown in Supplementary Table 3).

In our previous study [3] we showed that, at a cellular level, there was no significant correlation between the presence of sense RNA foci and the presence of sense RNA-derived poly-GA inclusions. In this study, we examined the relationship between the presence of antisense RNA foci and poly-PA inclusions in fifteen motor neurons from four *C9ORF72*-ALS cases (Supplementary Table 4). As for the sense species, there was no significant correlation between the two observations (χ^2, $p = 0.83$).

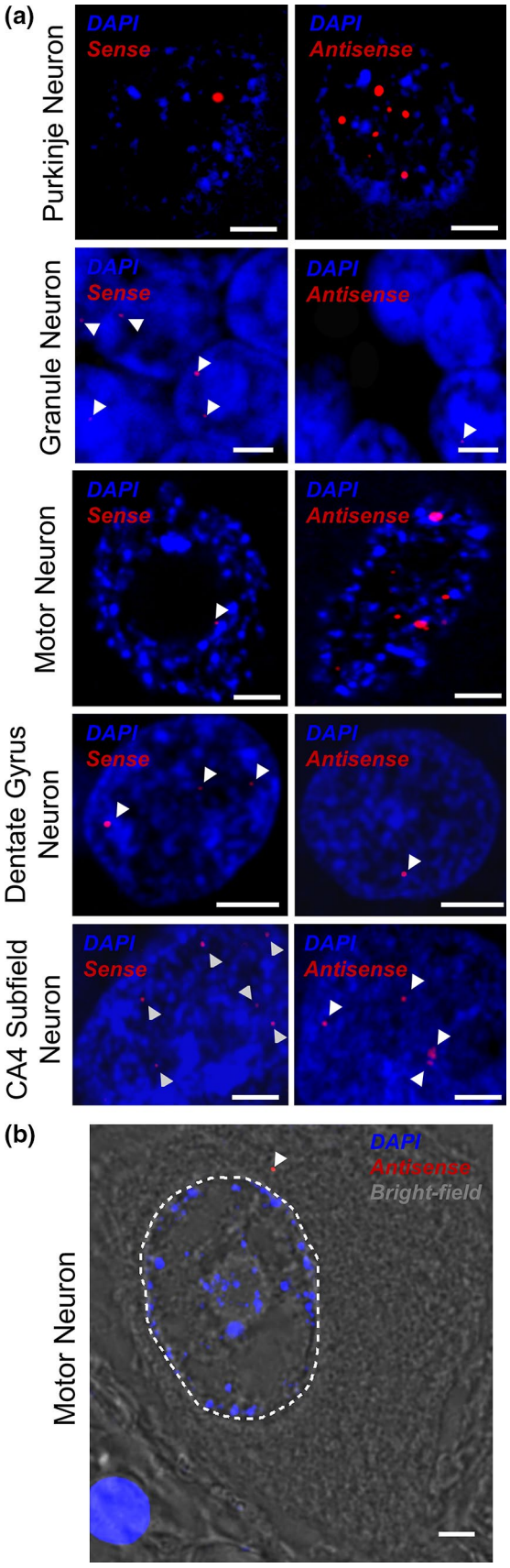


Fig. 1 RNA FISH reveals the distribution of sense and antisense RNA foci in five neuronal populations. Representative images show that antisense RNA foci are more numerous in cerebellar Purkinje neurons and motor neurons; in contrast sense, RNA foci are more numerous in cerebellar granule neurons; neither population is more abundant in dentate gyrus neurons and CA4 subfield neurons of the hippocampus (**a**). Smaller foci are highlighted by *arrowheads*. As has been previously demonstrated for sense foci, antisense foci are occasionally present in the cytoplasm of mature motor neurons (**b**, *arrowhead*, the nuclear border is indicated by a *dotted line*). *Scale bar* 3 µm

Cellular distribution of RNA foci and RRM-containing proteins

We used confocal microscopy to validate in vivo some of the RRM-containing proteins that were found to interact with $(CCCCGG)_4$ repeat RNA [11]. We and others have previously demonstrated colocalisation of SRSF2, hnRNP A1, hnRNP H/F, and ALYREF with sense RNA foci [3, 17]. We set out to investigate the cellular distribution of the same proteins with respect to antisense RNA foci, and we also examined nucleolin and hnRNP K which are proposed to be specific binding partners of sense and antisense foci, respectively [11].

Approximately 50 cerebellar Purkinje neurons were examined in a blinded experiment, from a minimum of three *C9ORF72*-ALS cases. Simultaneous co-staining was carried out in parallel in non-*C9ORF72* ALS cases and neurologically normal controls. For ALYREF, hnRNP A1, SRSF2, hnRNP H/F, and hnRNP K, the overall cellular distribution was not specifically altered in *C9ORF72*+ cases except for nuclear areas where colocalisation was demonstrated (Fig. 3a–e). Haeusler et al. [11] observed disruption of nucleolin expression from the nucleolus in cell models expressing expanded *C9ORF72,* but reported a variable distribution or nucleolin in *C9ORF72*+ CNS tissue. In agreement with this, we identified *C9ORF72*+ neurons which did and did not demonstrate an altered distribution pattern of nucleolin (Fig. 3f).

By IHC, we demonstrated colocalisation of SRSF2, hnRNP A1, hnRNP H/F, ALYREF, and hnRNP K in cerebellar Purkinje neurons with 34, 21, 3.4, 7.8, and 8.1 % of antisense RNA foci, respectively (Fig. 3a–e). In contrast, nucleolin was not observed to colocalise with antisense RNA foci (Fig. 3f). To validate the IHC findings, we performed in vitro Ultra-Violet (UV) crosslinking assays using radiolabelled synthetic $(GGGGCC)_5$ or $(CCCCGG)_5$ RNA oligonucleotides, and purified recombinant proteins synthesized in *E.coli*. Unlike IHC, this allows determination of direct and specific RNA:protein interactions via the formation of covalent bonds under UV light exposure. Both sense and antisense repeat RNA were observed to directly interact with hnRNP F, hnRNP A1, ALYREF, and

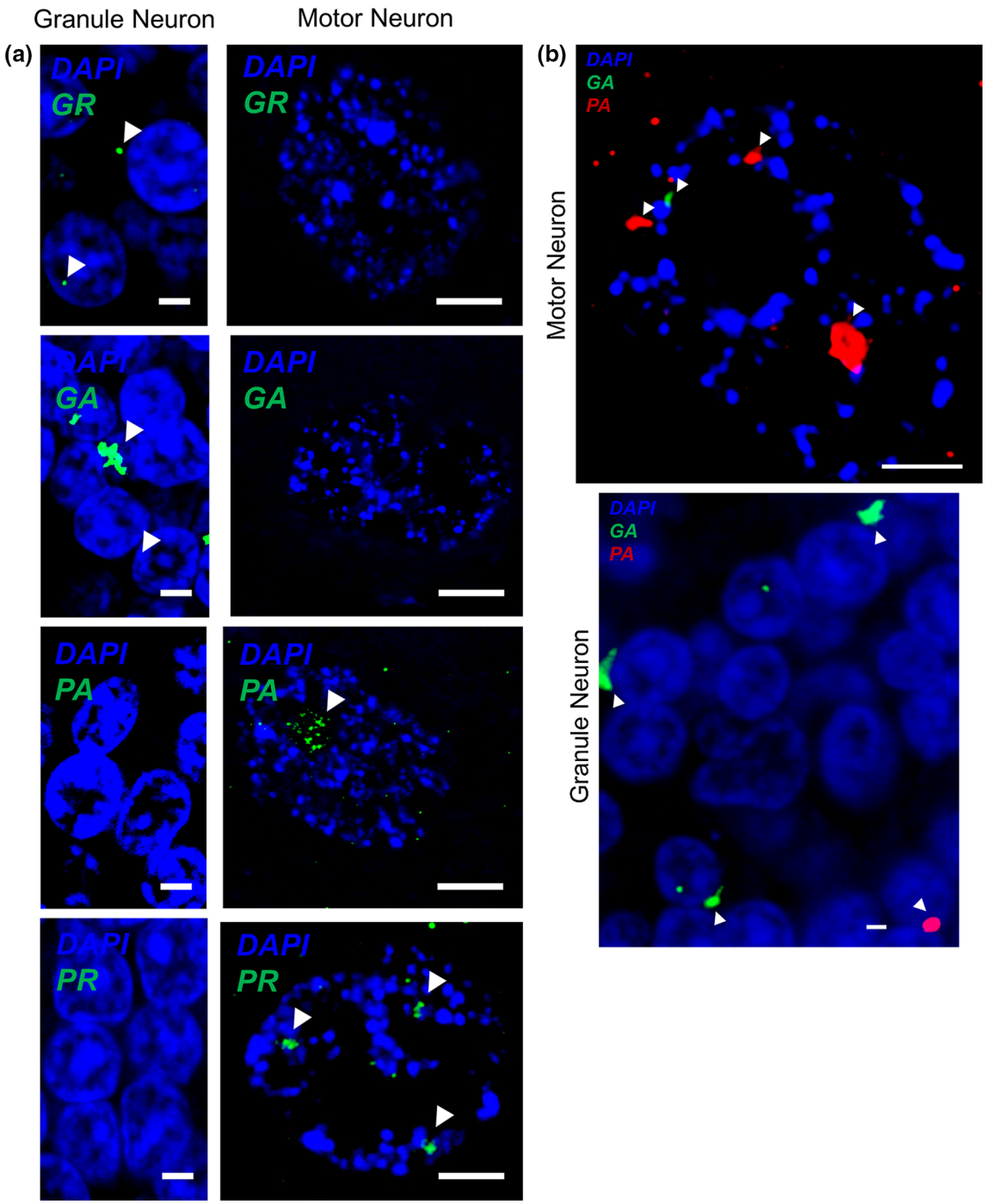


Fig. 2 Immunohistochemistry reveals the distribution of dipeptide repeat protein containing inclusions consisting of species derived from sense and antisense repeat RNAs in two neuronal populations. Representative images showing that poly-GA and poly-GR containing inclusions are more numerous in cerebellar granule neurons, whereas poly-PA and poly-PR containing inclusions are more numerous in motor neurons. Staining was carried out individually for each protein (**a**) and then poly-GA and poly-PA were examined by dual staining (**b**). Inclusions are highlighted by *arrowheads*. *Scale bar* 3 μm

SRSF2 proteins although the RNA-binding activity was not equal in all cases and for hnRNP A1 was relatively low (Fig. 4). In contrast, we failed to detect any direct interactions between sense or antisense repeat RNA and hnRNP K, suggesting that the previously observed colocalisation of hnRNP K with antisense RNA foci is not due to direct binding between hnRNP K and CCCCGG-repeats. The smeared appearance of certain of the proteins on the phospho image (Fig. 4) is likely to be due to the formation of covalently bonded oligomeric protein:RNA complexes.

Table 3 Mean and standard deviation (SD) of number of inclusions per cell containing poly-GA and poly-PA protein, in granule neurons and motor neurons from six patients with C9ORF72-disease

Case	Poly-PA (mean)	Poly-PA (SD)	Poly-GA (mean)	Poly-GA (SD)	*p* value
Motor neurons					
1	2.1	2.32	0	0	0
2	2.4	3.58	0.5	0.76	1.57E−34
4	2	1.83	0.3	0.35	0.0046
6	1.2	0.75	0	0	1.55E−21
7	1.7	1.56	0.2	0.40	0
8	4.4	2.07	0.2	0.44	0
Granule neurons					
1	0.04	0.18	0.2	0.36	2.42E−44
2	0.01	0.07	0.21	0.37	0
4	0.04	0.2	0.2	0.42	0
6	0.01	0.09	0.1	0.36	0
7	0.01	0.08	0.1	0.29	7.58E−49
8	0.01	0.12	0.1	0.34	0

In each case, poly-GA containing inclusions are significantly more numerous in granule neurons (likelihood-ratio test $p < 0.01$) and poly-PA containing inclusions are significantly more numerous in motor neurons (likelihood-ratio test $p < 0.01$)

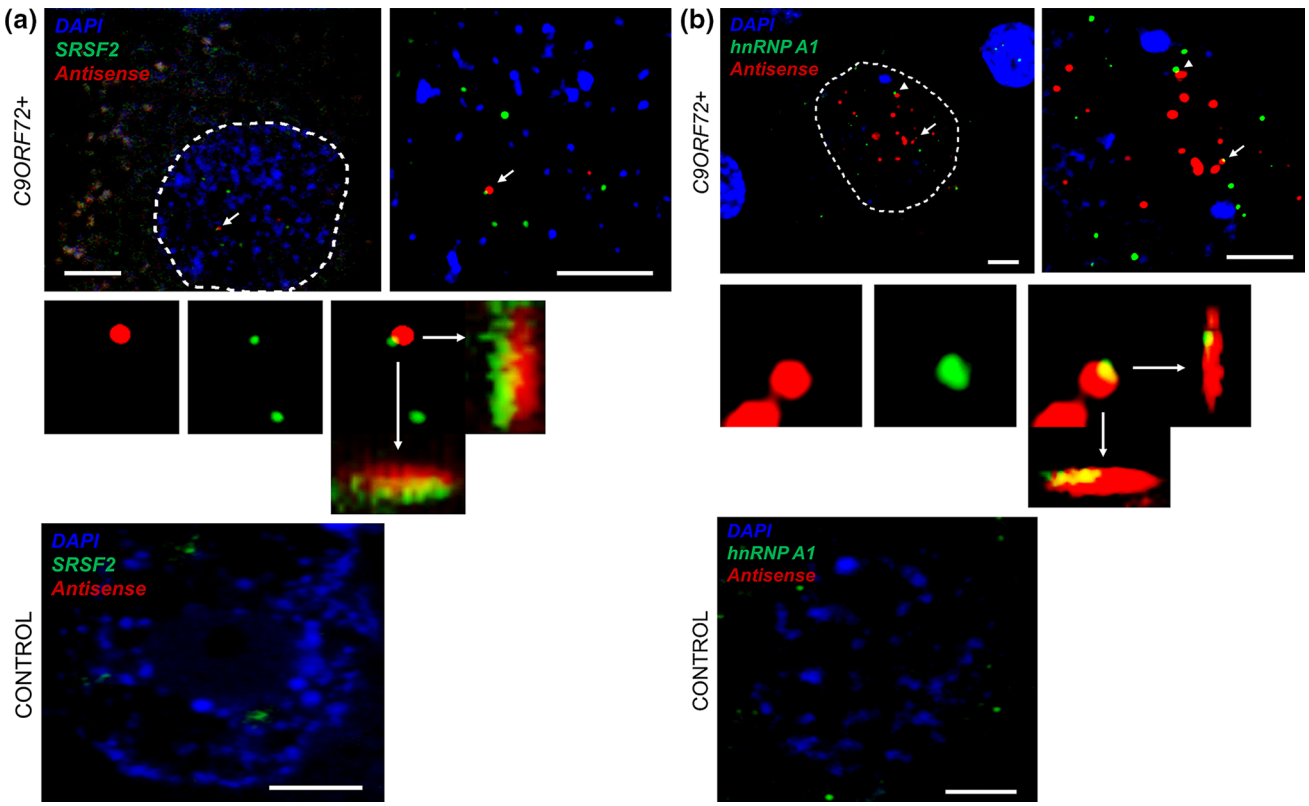


Fig. 3 Combined RNA FISH and IHC demonstrate colocalisation of nucleolin and nuclear speckle components with antisense RNA foci in Purkinje neurons from *C9ORF72*-ALS patients and the distribution of these proteins in Purkinje neurons from control individuals. SRSF2 (**a**), hnRNP A1 (**b**), hnRNP H/F (**c**), ALYREF (**d**), and hnRNP K (**e**) are observed to colocalise with antisense RNA foci (*arrows*) in Purkinje neurons from *C9ORF72*-ALS patients. A large scale view is shown to the *left* of a zoomed-in image. Colocalisation events are enlarged including orthogonal views, and unmerged protein and RNA foci are shown for comparison. There was not a significant difference between the staining of these proteins in controls and C9ORF72+ individuals, but no antisense RNA foci are observed in controls. Nucleolin was not observed to colocalise with antisense RNA foci (**f**); moreover, the distribution of nucleolin was variable in *C9ORF72*+ Purkinje neurons. In some cells, nucleolin was prominently nucleolar (**f**, *left panel*) and in other cells it was dispersed throughout the nucleus (**f**, *right panel*, RNA focus is indicated by an *arrowhead*). The *dotted line* illustrates the nuclear border in images **a**–**e** and the nucleolar border in image **f**. Scale bar 3 μm

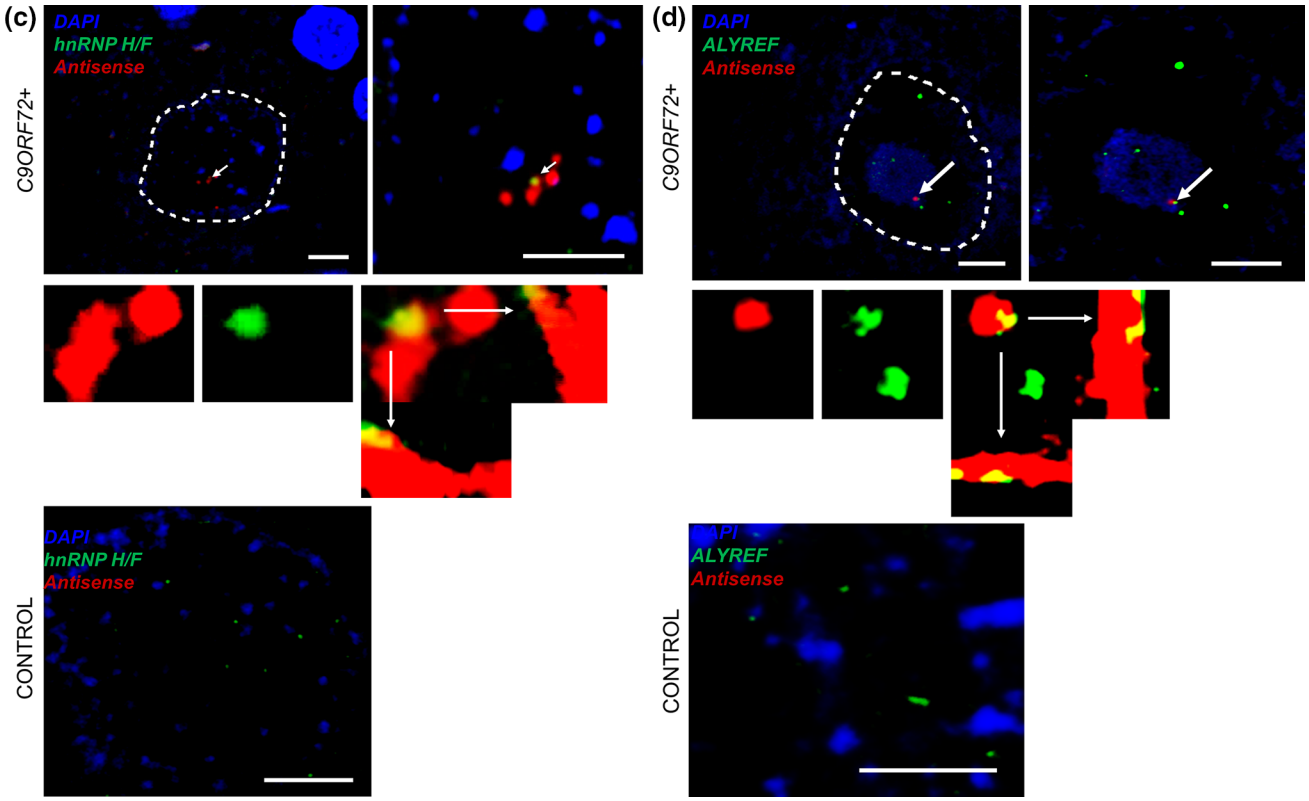


Fig. 3 continued

Multiple molecules of RRM-containing proteins bind to RNA oligonucleotides via inter and intra molecular interactions. We have demonstrated a similar effect previously [10].

Cellular distribution of RNA foci and TDP-43

We also examined the association of RNA foci with depletion of TDP-43 from the nuclei of motor neurons of seven patients with *C9ORF72*-ALS. Nuclear depletion and cytoplasmic mislocalisation of TDP-43 form the pathological hallmark of most subtypes of ALS, including *C9ORF72*-mediated disease [25]. We have previously shown that the proportion of sense RNA foci+ motor neurons with and without nuclear TDP-43 is approximately equivalent (χ^2, $p = 0.75$) [3]. As a direct comparison with this study, approximately fifty motor neurons were examined in FFPE sections from seven *C9ORF72*-ALS cases (Supplementary Table 5). Unlike sense RNA foci, the presence of antisense foci was significantly associated with nuclear loss of TDP-43. Seventy-seven percent of antisense foci+ motor neurons displayed loss of nuclear TDP-43 compared to 13 % of motor neurons without observable antisense foci (χ^2, $p < 0.00001$) (e.g. Fig. 5). A similar experiment in hippocampal CA4 subfield neurons did not reveal a significant correlation between the presence of antisense foci and nuclear loss of TDP-43, indeed no CA4 subfield neurons exhibited complete nuclear clearance of TDP-43 (data not shown).

Discussion

The precise mechanisms of neuronal injury in *C9ORF72*-disease appear complex, and are likely to involve RNA gain-of-function toxicity mediated by sense and antisense transcription of the GGGGCC repeat expansion (reviewed in [2, 4]). The small number of cases examined in this study and the wide variability in the phenotype of *C9ORF72*-related disease prohibits informative comparison between foci distribution and clinical phenotype, but a useful proxy is the pathological hallmark of ALS neurodegeneration: nuclear loss of TDP-43 [25]. We have demonstrated that antisense but not sense foci are significantly associated with nuclear loss of TDP-43 in motor neurons. This intriguing observation suggests that antisense RNA foci may occupy a key position in the cascade of disease pathogenesis. Moreover, examining the differences and similarities between the two species of RNA foci may shed light on important mechanisms leading to neurodegeneration.

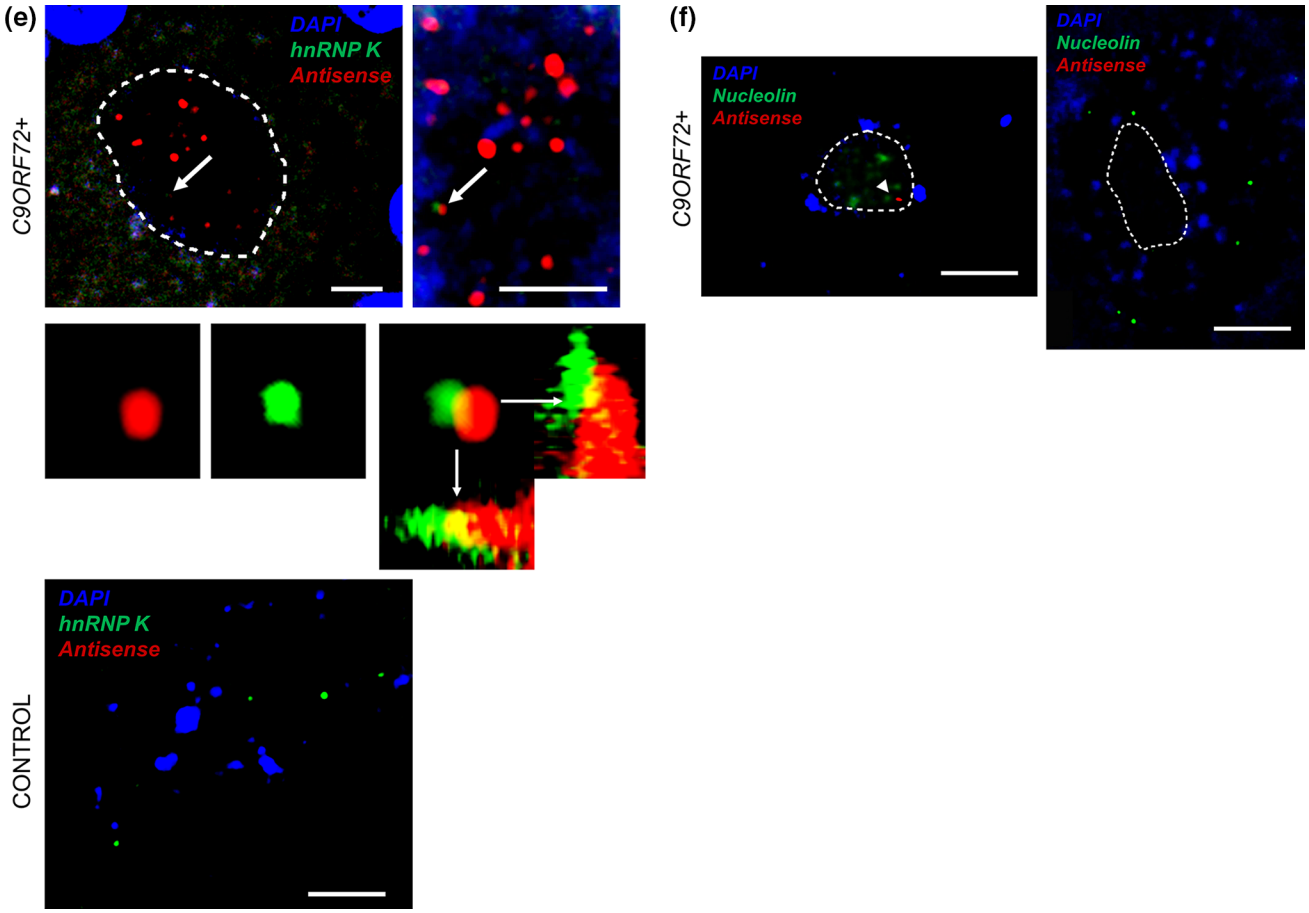


Fig. 3 continued

We have demonstrated colocalisation of antisense RNA foci with SRSF2, hnRNP A1, hnRNP H/F, ALYREF, and hnRNP K, but not nucleolin. This is consistent with the work of Haeusler et al. [11]. UV-crosslinking studies confirmed that each of these interactions is direct and specific, with the exception of hnRNP K. Conflicting results between the two methodologies may arise because IHC is unable to distinguish between direct and indirect interaction. There is significant potential for indirect binding: many RRM proteins co-exist and interact within nuclear speckles. Notably of the proteins we have examined, SRSF2 colocalisation with both sense [3] and antisense foci was observed with the highest frequency, and this protein is the core component of nuclear speckles [30].

Both the IHC and the UV-crosslinking studies in this report suggest that the binding partners of sense and antisense RNA foci are not significantly different. This is also reported by others [11]. Many of these identified binding partners are localised, with SRSF2, to nuclear speckles, nuclear domains implicated in the storage, and supply of splicing factors to active transcription sites [30]. Neuromuscular diseases, including type 1 myotonic dystrophy (DM1), have been associated with depletion of normal components of nuclear speckles [1, 29]. Sequestration of these proteins by sense or antisense RNA foci and consequent disruption of the normal function of these essential nuclear organelles might be a key event in the pathophysiology of *C9ORF72*-mediated neurodegeneration. If so, our results would predict that both species of RNA foci should be equally toxic. This is consistent with observed toxicity of sense foci in various model systems [8, 16, 17, 21, 28]. This led us to ask whether the key difference might not be in the interactions of the foci themselves, but in the neuronal populations in which sense and antisense foci are expressed.

In all cases, the relative frequency of sense and antisense foci varied consistently and significantly between neuronal populations. Importantly in motor neurons, the primary target of pathology in ALS, antisense foci are more abundant than sense foci. Therefore, we suggest that the key event determining toxicity leading to TDP-43 mislocalisation, of antisense as opposed to sense RNA foci, might be a propensity to produce antisense foci mediated by cell-specific transcriptional regulation. Alternatively sense RNA foci

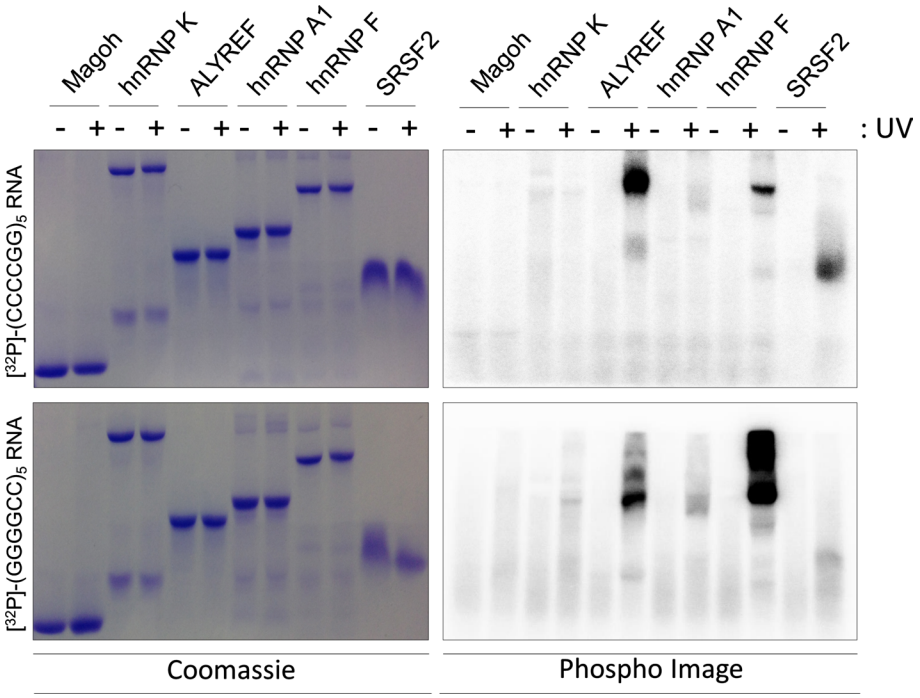


Fig. 4 Specific and direct interactions between $(GGGGCC)_5$ and/or $(CCCCGG)_5$ and hnRNP A1, hnRNP F, SFRS2, and ALYREF but not hnRNP K or Magoh (negative control). Magoh, SRSF2 9-101, ALYREF, hnRNP A1-like2, hnRNP K, and hnRNP F were expressed in *E. coli* and purified (see Supplementary Table 4). $(GGGGCC)_5$ (sense) and $(CCCCGG)_5$ (antisense) RNA oligonucleotides were end labelled with polynucleotide kinase using [ɣ-32P]-ATP, prior to incubation with purified proteins. RNA was covalently bound (+) or not (−) following UV irradiation. The absence of radioactive signal (*right panel*, PhosphoImage) in the absence of UV irradiation demonstrates specificity of direct binding observed after UV treatment. All gels shown in the *different panels* were exposed simultaneously for the same amount of time (4 h). Note that a high molecular weight band is also observed for ALYREF due to oligomerisation properties [10]

might be degraded at a higher rate than antisense RNA foci. *In-vitro* studies have suggested that both sense and antisense RNA sequences form complex secondary structures including G-quadruplexes and hairpin loops [11]. These secondary structures may help stabilize the RNA foci and prevent degradation.

It is interesting that two populations of relatively large neurons, motor neurons and cerebellar Purkinje neurons, exhibited antisense RNA foci at a higher frequency than sense RNA foci, in contrast to the smaller cerebellar granule neurons. This suggests that our observations may be related to some property correlated with neuronal size. However, in the hippocampus, neither the larger CA4 subfield neurons nor the smaller dentate granule neurons exhibited either species of RNA foci at a consistently higher frequency.

We observed antisense foci in the cytoplasm of motor neurons, which is consistent with aberrant nuclear export and may be a key step in the facilitation of proposed repeat associated non-ATG translation to produce DPR species [24]. We made a similar observation with respect to sense foci [3] and we suggest that interaction between repeat RNA and mRNA export adaptors, such as ALYREF, might override the normal nuclear retention of pre-mRNA species. Recent studies consistent with a key role for DPRs in the pathogenesis of *C9ORF72*-mediated neurodegeneration [15, 20, 22] suggest that this represents an attractive therapeutic target.

We have demonstrated that the frequency of sense and antisense foci is usually correlated i.e. a patient with more sense foci will also have more antisense foci. This is particularly interesting in case 1 from our analysis (Table 2) who displayed a relatively high frequency of sense and antisense RNA foci in the cerebellum and CA4 subfield neurons of the hippocampus, which are both extra-motor areas. Case 8 also exhibited a relatively high frequency of sense and antisense RNA foci in CA4 subfield neurons; the frequency of RNA foci in the cerebellum of case 8 was not quantified. In contrast to the other cases examined these patients displayed extra-motor disease clinically as well as pathologically: clinical FTD was diagnosed with (case 1) and without (case 8) ALS (Table 1). This is consistent with a correlation between the development of RNA foci in specific neuronal subtypes and clinical presentation, but this hypothesis will require validation in a larger number of FTD and ALS cases.

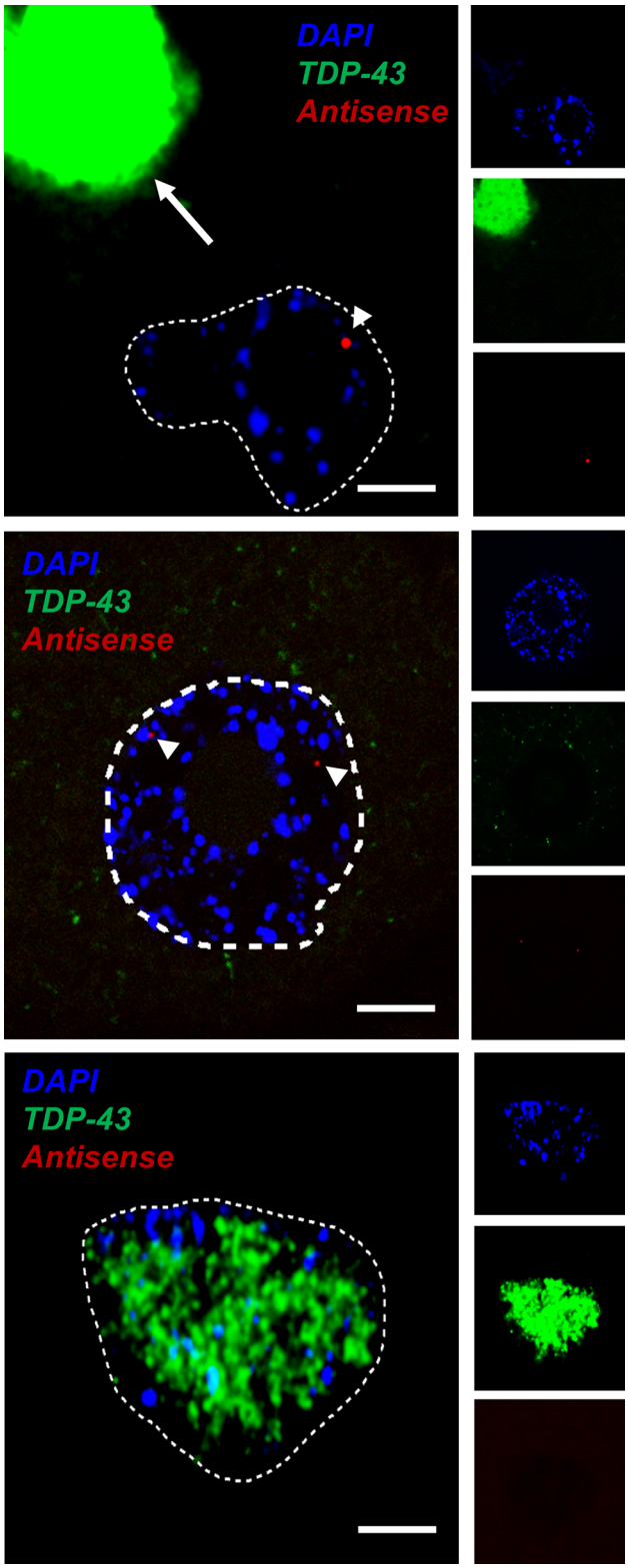


Fig. 5 TDP-43 IHC and RNA FISH demonstrate that antisense RNA ▶ foci are significantly associated with nuclear clearance of TDP-43 in motor neurons. Representative images showing that antisense RNA foci (*arrowheads*) are significantly associated with nuclear clearance of TDP-43 in motor neurons of *C9ORF72*-ALS patients; split channel images are provided for comparison. Cleared TDP-43 may be present within a cytoplasmic inclusion (*upper panels*; RNA focus is indicated by the *arrowhead*, a compact inclusion is *arrowed*) or simply present in the cytoplasm (*middle panels*; RNA foci are indicated by *arrowheads*). In contrast, the absence of antisense RNA foci is significantly associated with the presence of nuclear TDP-43 (*lower panels*). *Scale bar* 3 μm

Finally, varying frequency of the expression of sense and antisense repeat RNA has implications for the formation of specific DPRs. Our observations of all five DPRs are consistent with our conclusions relating to the expression of sense and antisense RNA foci. In cerebellar granule neurons, where sense RNA foci are more abundant, there is a higher frequency of sense-RNA derived DPR inclusions; and in motor neurons where antisense RNA foci are more abundant, there is a higher frequency of antisense RNA-derived DPR inclusions. Therefore, we suggest that, at least in these neuronal populations, translation of the sense and antisense derived proteins occurs in different quantities depending of the relative availability of RNA repeat molecules. The results of the present study potentially explain the observations of others that inclusions containing poly-GA protein are much more abundant in certain neuronal populations including cerebellar granule neurons [18]. However, Davidson et al. [6] failed to demonstrate antisense RNA derived DPR inclusions in Purkinje neurons of the cerebellum and dentate gyrus neurons of the hippocampus. This contrasts with our demonstration of antisense RNA foci in both of these populations, particularly in the cerebellar Purkinje neurons which we found to show a preference for exhibiting antisense rather than sense RNA foci. This variation between neuronal populations might be explained by variability in control of nuclear export of repeat RNA species; in this context, it is interesting that mutations in *hGle1*, a mRNA export adaptor, have recently been shown to cause selective death of motor neurons [14].

In our previous study we showed that there was no significant correlation between the presence or absence of nuclear sense RNA foci in cerebellar granule neurons and whether or not those cells contain a cytoplasmic inclusion positive for poly-GA [3]. Similarly in this study we have shown that there is no significant correlation between the presence or absence of nuclear antisense RNA foci in motor neurons, and whether or not those cells contain an inclusion positive for poly-PA. This suggests that our population level conclusion in this study, that neuronal populations have a propensity to produce either sense or antisense RNA derived foci and DPRs, does not apply at a cellular level. Thus, whilst both RNA foci and DPRs are derived from the same RNA molecules, the processes by which this RNA is stabilised into a focus or exported for translation are probably different or even mutually exclusive—indeed work from Gendron et al. [9] suggested that this might be the case. So if motor neurons have a preference for antisense transcription of the *C9ORF72* expansion then the

motor neuron population will express higher levels of antisense RNA derived foci and DPRs, but not necessarily within the same individual cells.

Our work highlights that any therapeutic approach to *C9ORF72*-ALS should consider the presence of antisense RNA foci in motor neurons. An antisense oligonucleotide approach has been proposed as a therapeutic option in *C9ORF72*-disease [8, 28]. We suggest that both sense and antisense RNAs should be targeted, as has been proposed by others [16]. Indeed, in relation to the relative selective vulnerability of motor neurons in *C9ORF72*-mediated pathology in vivo, targeting the antisense foci may be even more important than targeting sense foci. A limitation of our study is the reliance on post-mortem tissue which represents end-stage disease and may exclude the most vulnerable cells which have already been lost; as such we await validation of our findings in model systems.

Acknowledgments We acknowledge grants from the European Community's Seventh Framework Programme (FP7/2007-2013) under the EuroMOTOR project, Grant Agreement No 259867 and the EU Joint Programme—Neurodegenerative Disease Research (JPND) projects, SOPHIA and STRENGTH, supported through the following funding agencies under the aegis of JPND—www.jpmd.eu: United Kingdom, Medical Research Council to PJS and JK. PJS is an NIHR Senior Investigator. JCK and JRH are supported by MND Association/MRC Lady Edith Wolfson Fellowship awards ([MR/K003771/1] and [G0 800380] respectively). Biosample collection was supported by the MND Association and the Wellcome Trust (PJS). We are grateful to all of the patients with ALS and control individuals without neurological disease who donated biosamples for research.

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained for all individual participants included in the study.

References

1. Bengoechea R, Tapia O, Casafont I, Berciano J, Lafarga M, Berciano MT (2012) Nuclear speckles are involved in nuclear aggregation of PABPN1 and in the pathophysiology of oculopharyngeal muscular dystrophy. Neurobiol Dis 46:118–129
2. Cooper-Knock J, Shaw PJ, Kirby J (2014) The widening spectrum of *C9ORF72*-related disease; genotype/phenotype correlations and potential modifiers of clinical phenotype. Acta Neuropathol 127:333–345
3. Cooper-Knock J, Walsh MJ, Higginbottom A, Highley JR, Dickman MJ, Edbauer D et al (2014) Sequestration of multiple RNA Recognition Motif-containing proteins by *C9ORF72* repeat expansions. Brain 137:2040–2051
4. Cooper-Knock J, Kirby J, Highley R, Shaw PJ (2015) The spectrum of C9orf72-mediated neurodegeneration and amyotrophic lateral sclerosis. Neurotherapeutics. doi:10.1007/s13311-015-0342-1
5. Cruz-Migoni A, Hautbergue GM, Artymiuk PJ, Baker PJ, Bokori-Brown M, Chang CT et al (2011) A *Burkholderia pseudomallei* toxin inhibits helicase activity of translation factor eIF4A. Science 334:821–824
6. Davidson YS, Barker H, Robinson AC, Thompson JC, Harris J, Troakes C et al (2014) Brain distribution of dipeptide repeat proteins in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. Acta Neuropathol Commun 2:70
7. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ et al (2011) Expanded GGGGCC hexanucleotide repeat in noncoding region of *C9ORF72* causes chromosome 9p-linked FTD and ALS. Neuron 72:245–256
8. Donnelly CJ, Zhang PW, Pham JT, Heusler AR, Mistry NA, Vidensky S et al (2013) RNA toxicity from the ALS/FTD *C9ORF72* expansion is mitigated by antisense intervention. Neuron 80:415–428
9. Gendron TF, Bieniek KF, Zhang YJ, Jansen-West K, Ash PE, Caulfield T et al (2013) Antisense transcripts of the expanded *C9ORF72* hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. Acta Neuropathol 126:829–844
10. Golovanov AP, Hautbergue GM, Tintaru AM, Lian LY, Wilson SA (2006) The solution structure of REF2-I reveals interdomain interactions and regions involved in binding mRNA export factors and RNA. RNA 12:1933–1948
11. Haeusler AR, Donnelly CJ, Periz G, Simko EA, Shaw PG, Kim MS et al (2014) *C9ORF72* nucleotide repeat structures initiate molecular cascades of disease. Nature 507:195–200
12. Hautbergue GM, Hung ML, Golovanov AP, Lian LY, Wilson SA (2008) Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. Proc Natl Acad Sci 105:5154–5159
13. Hautbergue GM, Hung ML, Walsh MJ, Snijders AP, Chang CT, Jones R et al (2009) UIF, a new mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. Curr Biol 19:1918–1924
14. Kaneb HM, Folkmann AW, Belzil VV, Jao LE, Leblond CS, Girard SL et al (2015) Deleterious mutations in the essential mRNA metabolism factor, hGle1, in amyotrophic lateral sclerosis. Hum Mol Genet 24:1363–1373
15. Kwon I, Xiang S, Kato M, Wu L, Theodoropoulos P, Wang T et al (2014) Poly-dipeptides encoded by the *C9ORF72* repeats bind nucleoli, impede RNA biogenesis, and kill cells. Science 345:1139–1145
16. Lagier-Tourenne C, Baughn M, Rigo F, Sun S, Liu P, Li HR et al (2013) Targeted degradation of sense and antisense *C9ORF72* RNA foci as therapy for ALS and frontotemporal degeneration. Proc Natl Acad Sci 110:E4530–E4539
17. Lee YB, Chen HJ, Peres JN, Gomez-Deza J, Attig J, Stalekar M et al (2013) Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. Cell Rep 5:1178–1186
18. Mackenzie IR, Arzberger T, Kremmer E, Troost D, Lorenzl S, Mori K et al (2013) Dipeptide repeat protein pathology in *C9ORF72* mutation cases: clinico-pathological correlations. Acta Neuropathol 126:859–879

19. Mackenzie IR, Frick P, Neumann M (2014) The neuropathology associated with repeat expansions in the *C9ORF72* gene. Acta Neuropathol 127:347–357
20. May S, Hornburg D, Schludi MH, Arzberger T, Rentzsch K, Schwenk BM et al (2014) *C9ORF72* FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. Acta Neuropathol 128:485–503
21. Mizielinska S, Lashley T, Norona FE, Clayton EL, Ridler CE, Fratta P et al (2013) *C9ORF72* frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. Acta Neuropathol 126:845–857
22. Mizielinska S, Gronke S, Niccoli T, Ridler CE, Clayton EL, Devoy A et al (2014) *C9ORF72* repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. Science 345:1192–1194
23. Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E et al (2013) The *C9ORF72* GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science 339:1335–1338
24. Mori K, Arzberger T, Grasser FA, Gijselinck I, May S, Rentzsch K et al (2013) Bidirectional transcripts of the expanded *C9ORF72* hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. Acta Neuropathol 126:881–893
25. Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT et al (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science 314:130–133
26. Phelan MM, Goult BT, Clayton JC, Hautbergue GM, Wilson SA, Lian LY (2012) The structure and selectivity of the SR protein SRSF2 RRM domain with RNA. Nucleic Acids Res 40:3232–3244
27. Renton AE, Majounie E, Waite A, Simon-Sanchez J, Rollinson S, Gibbs JR et al (2011) A hexanucleotide repeat expansion in *C9ORF72* is the cause of chromosome 9p21-linked ALS-FTD. Neuron 72:257–268
28. Sareen D, O'Rourke JG, Meera P, Muhammad AK, Grant S, Simpkinson M et al (2013) Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a *C9ORF72* repeat expansion. Sci Transl Med 5:208ra149
29. Smith KP, Byron M, Johnson C, Xing Y, Lawrence JB (2007) Defining early steps in mRNA transport: mutant mRNA in myotonic dystrophy type I is blocked at entry into SC-35 domains. J Cell Biol 178:951–964
30. Spector DL, Lamond AI (2011) Nuclear speckles. Cold Spring Harb Perspect Biol 3:a000646

PLOS | ONE

RESEARCH ARTICLE

C9ORF72 GGGGCC Expanded Repeats Produce Splicing Dysregulation which Correlates with Disease Severity in Amyotrophic Lateral Sclerosis

Johnathan Cooper-Knock[1], Joanna J. Bury[1], Paul R Heath[1], Matthew Wyles[1], Adrian Higginbottom[1], Catherine Gelsthorpe[1], J. Robin Highley[1], Guillaume Hautbergue[1], Magnus Rattray[2], Janine Kirby[1], Pamela J. Shaw[1]*

1 Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, 385A Glossop Road, Sheffield, S10 2HQ, United Kingdom, **2** Life Sciences, The University of Manchester, Michael Smith Building, Oxford Road, Manchester, M13 9PT, United Kingdom

* pamela.shaw@sheffield.ac.uk

OPEN ACCESS

Citation: Cooper-Knock J, Bury JJ, Heath PR, Wyles M, Higginbottom A, Gelsthorpe C, et al. (2015) *C9ORF72* GGGGCC Expanded Repeats Produce Splicing Dysregulation which Correlates with Disease Severity in Amyotrophic Lateral Sclerosis. PLoS ONE 10(5): e0127376. doi:10.1371/journal.pone.0127376

Academic Editor: Huaibin Cai, National Institute of Health, UNITED STATES

Received: January 4, 2015

Accepted: April 15, 2015

Published: May 27, 2015

Data Availability Statement: The microarray dataset is available on GEO, accession number GSE68608. All relevant data are within the paper and its Supporting Information files.

Funding: The authors acknowledge grants from the 'EU Framework 7' (Euromotor No259867) and the SOPHIA project (funded by 'EU Joint Programme - Neurodegenerative Disease Research' and 'Medical Research Council') to PJS and JK. PJS is an NIHR Senior Investigator. JCK and JRH are supported by 'Motor Neurone Disease Association' / 'Medical Research Council' Lady Edith Wolfson Fellowship

Abstract

Objective

An intronic GGGGCC-repeat expansion of *C9ORF72* is the most common genetic variant of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia. The mechanism of neurodegeneration is unknown, but a direct effect on RNA processing mediated by RNA foci transcribed from the repeat sequence has been proposed.

Methods

Gene expression profiling utilised total RNA extracted from motor neurons and lymphoblastoid cell lines derived from human ALS patients, including those with an expansion of *C9ORF72*, and controls. In lymphoblastoid cell lines, expansion length and the frequency of sense and antisense RNA foci was also examined.

Results

Gene level analysis revealed a number of differentially expressed networks and both cell types exhibited dysregulation of a network functionally enriched for genes encoding 'RNA splicing' proteins. There was a significant overlap of these genes with an independently generated list of GGGGCC-repeat protein binding partners. At the exon level, in lymphoblastoid cells derived from *C9ORF72*-ALS patients splicing consistency was lower than in lines derived from non-C9ORF72 ALS patients or controls; furthermore splicing consistency was lower in samples derived from patients with faster disease progression. Frequency of sense RNA foci showed a trend towards being higher in lymphoblastoid cells derived from patients with shorter survival, but there was no detectable correlation between disease severity and DNA expansion length.

awards ([MR/K003771/1] and [G0 800380] respectively). Samples used in this research were in part obtained from the UK MND DNA Bank for MND Research, funded by the 'Motor Neurone Disease Association' and the 'Wellcome Trust'. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Significance

Up-regulation of genes encoding predicted binding partners of the *C9ORF72* expansion is consistent with an attempted compensation for sequestration of these proteins. A number of studies have analysed changes in the transcriptome caused by *C9ORF72* expansion, but to date findings have been inconsistent. As a potential explanation we suggest that dynamic sequestration of RNA processing proteins by RNA foci might lead to a loss of splicing consistency; indeed in our samples measurement of splicing consistency correlates with disease severity.

Introduction

GGGGCC repeat expansions within intron 1 of the *C9ORF72* gene are the most common cause of familial amyotrophic lateral sclerosis (ALS) and familial frontotemporal degeneration (FTD) [1,2], though how this genetic change results in neuronal injury is not yet understood. Evidence is being gathered for a gain-of-function toxicity mediated by either sequestration of RNA binding proteins (RBPs) by RNA foci transcribed from the repeat sequence [3–8], or via repeat associated non-ATG (RAN) translation of the repeat sequence to produce a dipeptide repeat protein [9–11], or a combination of both mechanisms.

Gene expression profiling has the potential to identify biological pathways aberrantly affected by the *C9ORF72* expansion. In addition, if toxicity is mediated by nuclear RNA foci developed from an intronic expansion, then transcriptome changes may be relatively upstream in disease pathogenesis [12]. On this basis we have studied gene expression changes in motor neurons and lymphoblastoid cell lines derived from individuals with *C9ORF72*-ALS.

We have previously suggested that dynamic sequestration by RNA foci of a number of RBPs might affect nuclear speckle function and thus disrupt mRNA splicing [8]. It has been proposed that splicing errors are a normal occurrence for which the cell is able to compensate [13]. Therefore an excessive splicing error rate may not immediately result in disease; however in time compensatory mechanisms might be overwhelmed in vulnerable cells. This is more consistent with the variable phenotype and late age of onset seen in *C9ORF72*-ALS than a model of binary toxicity resulting from a small number of specific splicing errors. Therefore we aimed to derive a measure of the overall splicing error rate in biosamples containing the *C9ORF72* repeat expansion. Additionally we used Southern hybridisation and RNA fluorescence in-situ hybridisation (FISH) to examine the relationship between the changes in the splicing error rate, disease severity, the length of the GGGGCC repeat expansion and the abundance of RNA foci.

Results

Transcriptome analysis

Motor neurons. Network analysis using WGCNA identified six significant networks within 5,000 genes considered (**Fig 1**) all of which were differentially expressed between *C9ORF72*-ALS and control groups, and showed significant functional enrichment (**Table 1**). Based on the median fold change, three networks were down-regulated and three networks were up-regulated in *C9ORF72*-ALS (**Table 1**). Specifically, within the brown network which was significantly enriched for transcripts related to the Gene Ontology (GO) term 'RNA splicing', 58.2% of transcripts were up-regulated. The yellow and green networks were also up-regulated and functionally enriched for 'male sex differentiation' and 'erythrocyte homeostasis' respectively. The

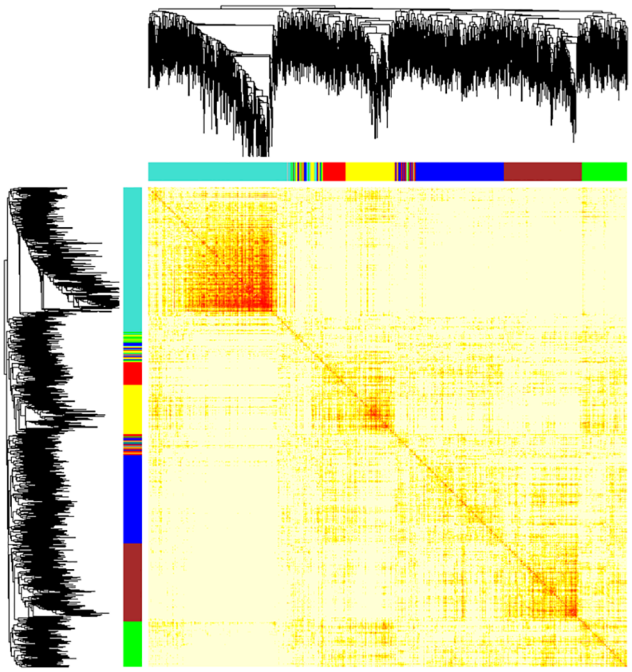


Fig 1. Gene level network analysis of transcriptome changes in motor neurons from *C9ORF72*-ALS cases. WGCNA analysis identified six gene networks which were dysregulated between *C9ORF72*-ALS and control samples. A clustering tree and heat map are shown illustrating separation of the gene networks, a lower branch height or darker colour denotes a greater Pearson correlation coefficient between pairs of genes.

doi:10.1371/journal.pone.0127376.g001

turquoise, blue and red networks were all down-regulated and functionally enriched for 'cholesterol biosynthetic process', 'regulation of glucose metabolic process' and 'regulation of nuclear division' respectively.

Lymphoblastoid cell lines. Two samples failed the Affymetrix quality control (QC) assessment and were excluded from the analysis based upon a low % presence call and/or AUC value. Network analysis using WGCNA identified nine significant networks which were differentially expressed between *C9ORF72*-ALS and control groups (Fig 2), and showed significant functional enrichment (Table 2). Based on the median fold change, five networks were down-

Table 1. Gene level network analysis of transcriptome changes in motor neurons from *C9ORF72*-ALS cases.

Network	Number of Genes	P-value C9ORF72-ALS Vs Control	Top Gene Ontology Enrichment	P-value for Enrichment Analysis	Median Fold Change
Turquoise	1555	0.008	Cholesterol biosynthetic process	0.001	0.62
Blue	1020	0.003	Regulation of glucose metabolic process	0.01	0.47
Brown	901	0.008	RNA splicing	7.45E-04	1.49
Yellow	635	0.003	Male sex differentiation	0.02	1.91
Green	579	0.0005	Erythrocyte homeostasis	0.01	1.75
Red	321	0.006	Regulation of nuclear division	0.01	0.49

WGCNA analysis identified six gene networks which were dysregulated between *C9ORF72*-ALS and control samples. The median fold change of genes within each network and the functional enrichment of each of the gene networks is tabulated. A fold change of >1 equates to up-regulation and a fold change of <1 equates to down-regulation.

doi:10.1371/journal.pone.0127376.t001

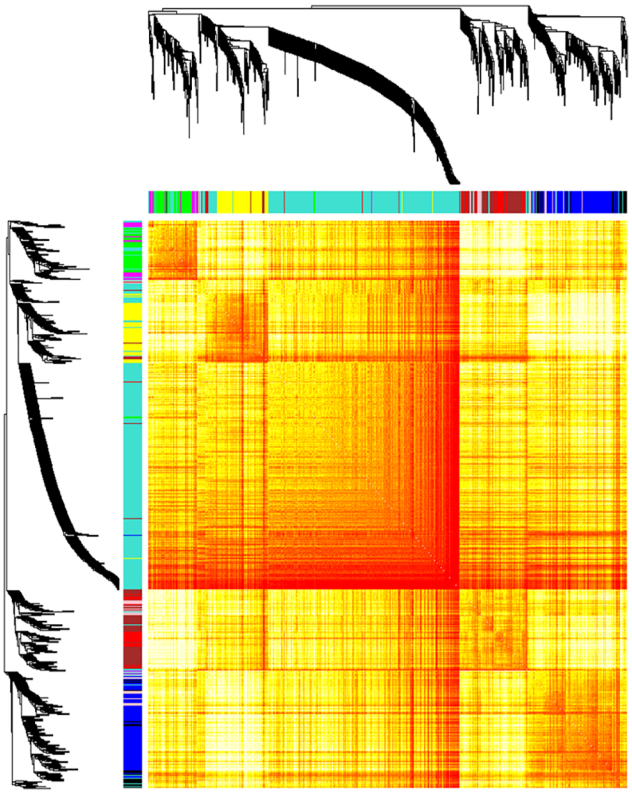


Fig 2. Gene level network analysis of transcriptome changes in lymphoblastoid cell lines derived from *C9ORF72*-ALS cases. WGCNA analysis identified nine gene networks which were dysregulated between *C9ORF72*-ALS and control samples. A clustering tree and heat map are shown illustrating separation of the gene networks, a lower branch height or darker colour denotes a greater Pearson correlation coefficient between pairs of genes. The median fold change of genes within each network and the functional enrichment of each of the gene networks is tabulated (B). A fold change of >1 equates to up-regulation and a fold change of <1 equates to down-regulation.

doi:10.1371/journal.pone.0127376.g002

Table 2. Gene level network analysis of transcriptome changes in lymphoblastoid cell lines derived from C9ORF72-ALS cases.

Network	Number of Genes	P-value C9ORF72-ALS Vs Control	Top Gene Ontology Enrichment	P-value for Enrichment Analysis	Median Fold Change
Turquoise	4653	8.11E-64	Positive regulation of apoptosis	0.01	1.19
Blue	1403	0.0000001	Regulation of action potential in neuron	0.02	0.88
Brown	1038	4.69E-09	Protein catabolic process	0.002	0.79
Yellow	854	2.48E-09	Synaptic transmission	0.004	1.16
Green	537	0.0000002	RNA splicing	1.50E-05	1.27
Red	427	9.17E-08	Positive regulation of apoptosis	0.02	0.74
Black	391	3.43E-09	Striated muscle tissue development	0.02	0.86
Pink	367	3.54E-08	Inflammatory response	0.004	0.86
Magenta	336	0.0000001	Protein catabolic process	1.43E-05	1.53

WGCNA analysis identified nine gene networks which were dysregulated between C9ORF72-ALS and control samples. The median fold change of genes within each network and the functional enrichment of each of the gene networks is tabulated. A fold change of >1 equates to up-regulation and a fold change of <1 equates to down-regulation.

doi:10.1371/journal.pone.0127376.t002

regulated and four networks were up-regulated in *C9ORF72*-ALS (Table 2). Specifically 92% of transcripts were up-regulated within the green network, which was significantly enriched for transcripts related to the GO term 'RNA splicing'. The turquoise and red networks were both functionally enriched for genes related to 'positive regulation of apoptosis', the blue and yellow networks were functionally enriched for categories directly related to neuronal function, and the black network was functionally enriched for 'striated muscle tissue development'; the pink network was functionally enriched for 'inflammatory response'; and the brown and magenta networks were functionally enriched for 'protein catabolic process.'

Analysis of networks enriched for the GO term 'RNA Splicing'. The brown network in the motor neurons and the green network in the lymphoblastoid cell lines were both up-regulated in *C9ORF72*-ALS samples and significantly enriched for transcripts related to the GO term 'RNA splicing.' We set out to determine whether the two networks contained similar transcripts or only transcripts with similar functional enrichment.

To make this comparison we reverted to the original lists of transcripts not filtered by Pearson correlation coefficient because computational burden was no longer an issue, and we were interested in all transcripts associated with the network signal and not just the most correlated. Examination of all genes significantly correlated (as quantified by Pearson correlation coefficient) ($p<0.05$) with the brown network signal and associated with the GO term 'RNA Processing' in motor neurons derived from *C9ORF72*-ALS patients revealed 88 transcripts encoding 74 unique genes (S1 Table). Examination of all genes significantly correlated ($p<0.05$) with the green network signal and associated with the GO term 'RNA Processing' in lymphoblastoid cells derived from *C9ORF72*-ALS patients revealed 459 transcripts encoding 236 unique genes (S1 Table). Given the difference in cell types and microarray platforms, there was evidence for significant similarity between the lists: 54% of the motor neuron list was also present within the lymphoblastoid cell list. Previously we have identified candidate binding partners of the GGGGCC repeat expansion by RNA pulldown and mass spectroscopy [8]. 20% of the unique hits identified in this way were present within the lymphoblastoid cell line list ($p<0.0001$) of which 89% were up-regulated in the *C9ORF72*-ALS samples compared to controls (S1 Table); and 10% of the unique hits were present within the motor neuron list ($p<0.0001$) of which 77% were up-regulated in the *C9ORF72*-ALS samples compared to controls (S1 Table).

Analysis of splicing. There was no significant difference in the total number of splicing events observed in lymphoblastoid cell lines derived from *C9ORF72*-ALS patients, non-*C9ORF72*-ALS patients and controls (Fig 3). However, the nature of those splicing events was significantly different. It is expected that functionally appropriate splicing would be similar in samples of a particular group and therefore we propose that splicing consistency is a marker of the error rate in RNA splicing. Splicing consistency was significantly reduced in the *C9ORF72*-ALS group compared to non-*C9ORF72*-ALS patients and controls (Fig 4A, S1 Fig). It is noteworthy that control cases do not share a common disease process and therefore might be expected to have quite different patterns of splicing. This is good evidence that splicing in *C9ORF72-ALS* is actively disrupted. In addition, splicing was less consistent in *C9ORF72-ALS* patients who lived <2 years following diagnosis compared to those that lived >4 years suggesting a link with the aggressiveness of the disease course (Fig 4B).

qPCR based validation of transcriptome changes

Candidates for qPCR validation were chosen from those genes which were up-regulated in the lymphoblastoid cells derived from *C9ORF72*-ALS cases compared to controls, and also identified as candidate binding partners of the GGGGCC repeat expansion by RNA pulldown and mass

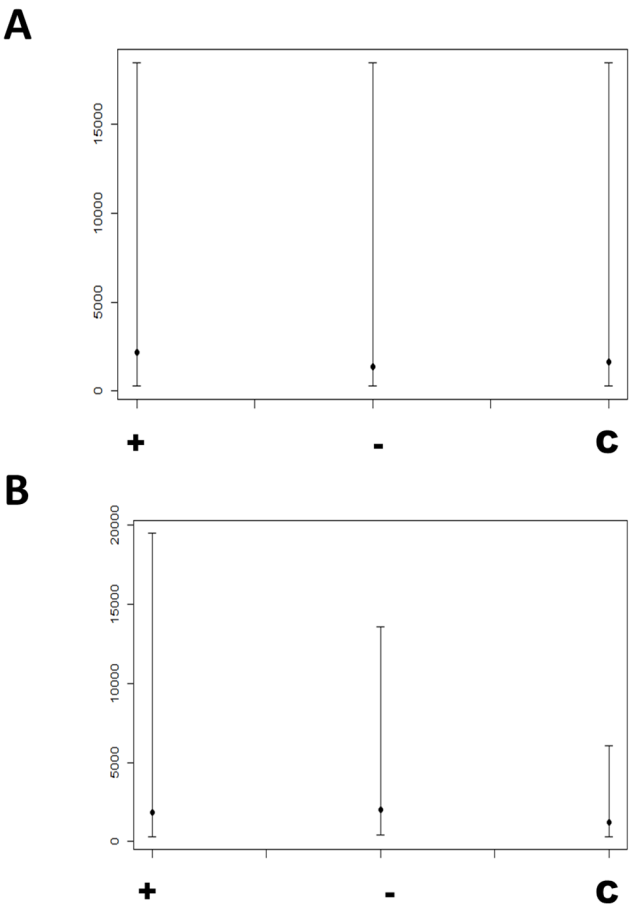


Fig 3. Frequency of exon inclusion and exclusion events. Plots of median and 95% CI for numbers of (A) exon inclusion and (B) exon exclusion events in *C9ORF72*-ALS (+), non-*C9ORF72* ALS (-) and control (C) derived lymphoblastoid cell lines, as determined by FIRMA score. There was no significant difference between sample groups.

doi:10.1371/journal.pone.0127376.g003

spectroscopy [8]. qPCR confirmed up-regulation of *HNRNPF* (1.43 fold, t-test, p = 0.001), *RBM3* (1.18 fold, t-test, p = 0.03) and *FUS* (1.35 fold, t-test, p = 0.005) but not *HNRNPH2*.

Estimation of expansion size and quantification of the abundance of RNA foci in lymphoblastoid cell lines

GGGGCC repeat expansion size and abundance of sense and antisense RNA foci was determined in lymphoblastoid cell lines derived from 17 patients with short (<2 years) disease duration and 7 patients with long (>4 years) disease duration. No difference in minimum (t-test, p = 0.10), modal (t-test, p = 0.41) or maximum (t-test, p = 0.57) repeat size was detectable between groups by Southern blotting (data not shown).

The frequency of lymphoblastoid cells containing sense RNA foci was higher in lines derived from 3 patients with short (<2 years) disease duration compared to lines derived from 3 patients with long (>4 years) disease duration, however this trend did not reach significance (average frequency of sense foci+ cells was 0.35 versus 0.12, t-test, p = 0.099). There was no such trend in the frequency of lymphoblastoid cells containing antisense RNA foci (average frequency of antisense foci+ cells 0.20 versus 0.32, t-test, p = 0.29). Example cells are shown in Fig 5.

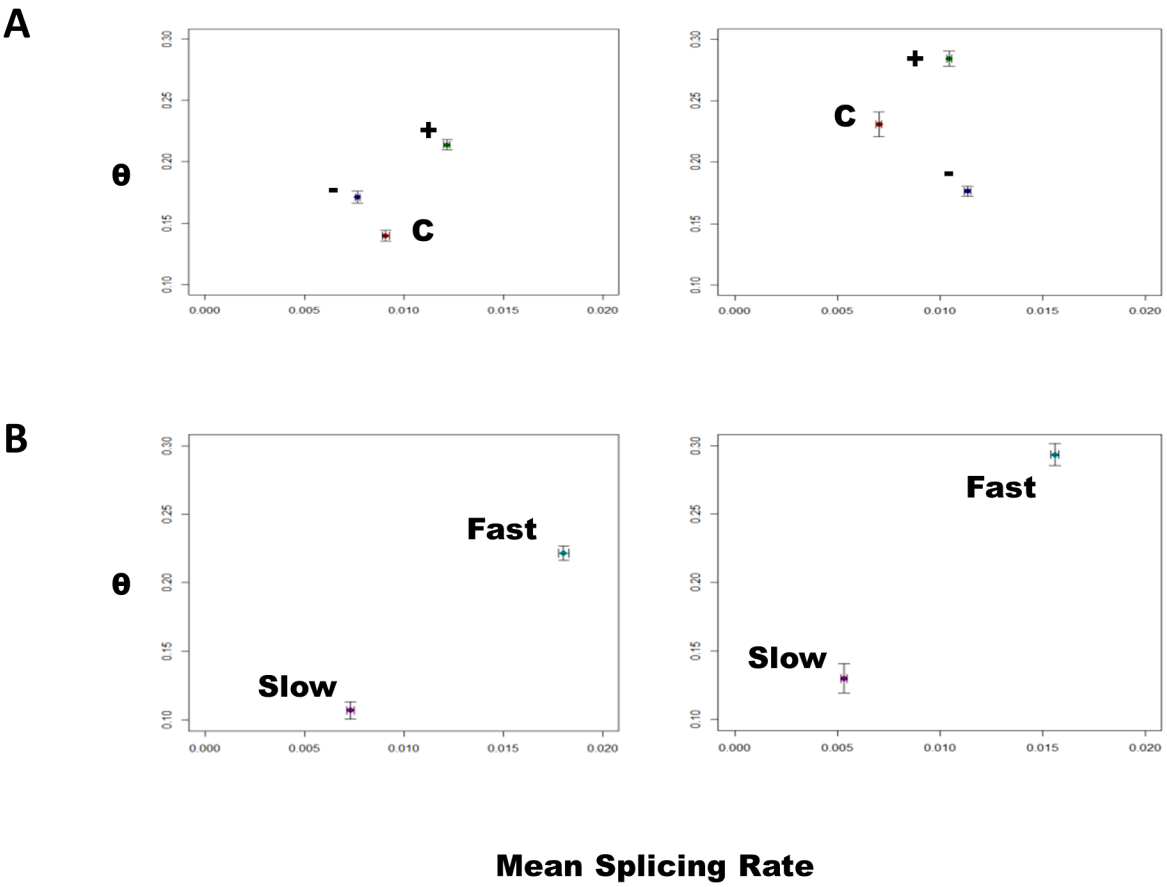


Fig 4. Plots of θ against the mean splicing rate with 95% confidence intervals. Exon inclusion events are shown in the left panel and exclusion inclusion events are shown in the right panel. θ is higher indicating reduced consistency of splicing in (A) *C9ORF72-ALS* (+) compared to non-*C9ORF72* ALS (-) and control (C) derived lymphoblastoid cell lines; and (B) in cell lines derived from patients with rapid (length <2 years, Fast) compared to slowly (length >4 years, Slow) progressive *C9ORF72*-ALS.

doi:10.1371/journal.pone.0127376.g004

Discussion

There is an urgent need to understand the mechanisms of neuronal injury in *C9ORF72*-disease. In order to establish the biological pathways altered by the presence of the GGGGCC repeat expansion we carried out gene expression profiling of isolated motor neurons from spinal cord and lymphoblastoid cell lines derived from human ALS patients and controls. Moreover, it has been suggested that the *C9ORF72* expansion has a direct effect on the transcriptome, possibly via the formation of RNA foci [1,8]; if this is the case then transcriptome changes may represent a relatively upstream component of pathogenesis and a suitable therapeutic target.

Transcriptome analysis in *C9ORF72*-ALS motor neurons

Six gene networks were identified as differentially expressed between *C9ORF72-ALS* and control motor neurons (**Fig 1**, **Table 1**). Several networks were significantly enriched for GO categories previously implicated in ALS including 'cholesterol biosynthetic process' which occurs primarily in the endoplasmic reticulum (ER) [14], 'regulation of glucose metabolic process' [15], 'regulation of nuclear division' [16] and 'RNA splicing' [17] (**Table 1**). The 'RNA splicing' network overlapped with a similar network in the lymphoblastoid cells and will be discussed further below.

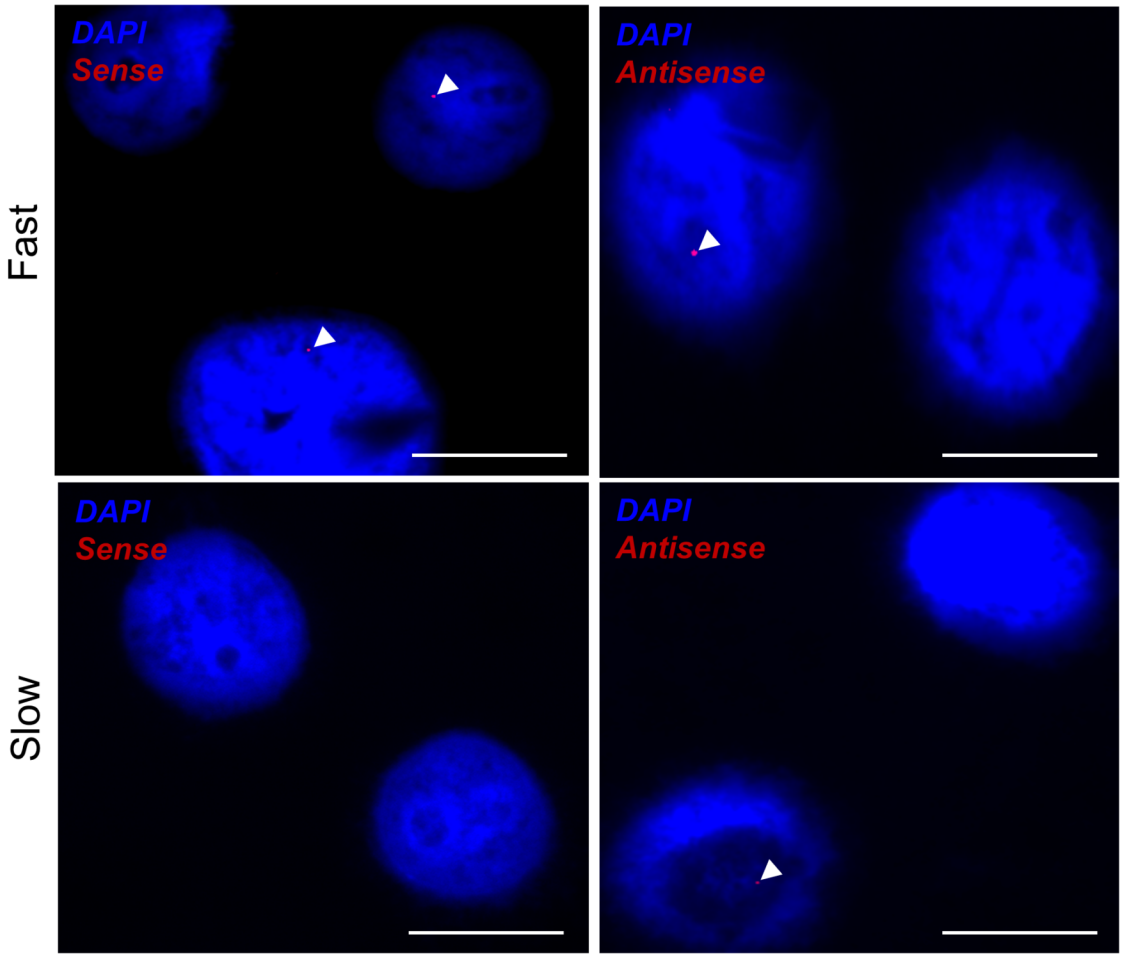


Fig 5. RNA foci in lymphoblastoid cell lines derived from patients with short or long survival. RNA FISH was performed for sense and antisense RNA foci in lymphoblastoid cells. Example cells are shown derived from patients with rapid (length <2 years, Fast, upper panels) compared to slowly (length >4 years, Slow, lower panels) progressive *C9ORF72*-ALS. GGGGCC-repeat sense RNA foci are visualised (arrowheads) in the left panels whereas GGCCCC-repeat antisense RNA foci are visualised (arrowheads) in the right panels. Scale bar 10 μm.

doi:10.1371/journal.pone.0127376.g005

Dysregulation of gene networks related to glucose and cholesterol metabolism, both of which were down-regulated in *C9ORF72*-ALS motor neurons, is interesting. Increasingly ER stress is implicated in ALS. As well as a site of lipid synthesis, the ER is responsible for correct protein folding [18]. Protein aggregates are a prominent feature of all forms of ALS; ER stress activates the unfolded protein response (UPR) and chronically can lead to apoptosis. Indeed activation of the UPR has been observed in sporadic ALS patients [19]. ER stress has been observed to impact negatively on cholesterol synthesis [20]. In a different cell model we have previously demonstrated a deficit in glucose metabolism associated with ALS [21].

Transcriptome analysis in *C9ORF72*-ALS lymphoblastoid cells

Nine transcript networks were identified as differentially expressed between *C9ORF72*-ALS and control lymphoblastoid cells (**Fig 2**, **Table 2**). With the exception of the gene network enriched for 'RNA splicing', the functional enrichment of differentially expressed networks in the *C9ORF72*-ALS lymphoblastoid cells and motor neurons was distinct. This is not unexpected given the use of non-overlapping cases in the sample sets and different analysis platforms;

indeed this makes the identification of a common network all the more robust. Moreover, there are differences in physiology between the cell types: motor neurons are post-mitotic whereas lymphoblastoid cells are actively dividing which may explain why 'regulation of nuclear division' did not appear in the lymphoblastoid cell networks.

Some common themes arise in the enrichment of the differentially expressed networks in the lymphoblastoid cells: three networks were enriched for functional categories related to nerve or muscle function (**Table 2**). This suggests that the presence of the expansion in the lymphoblastoid cells has an effect on genes important for neuromuscular function. Even if not deleterious to the lymphoblastoid cells, these same changes may be toxic if they occur in the cell types vulnerable to the neurodegenerative pathology in ALS. Two networks were enriched for functional categories related to protein catabolism (**Table 2**). Failure of protein catabolism has been implicated previously in ALS [22], indeed several genetic variants of ALS are caused by mutations in genes with roles in protein degradation e.g. VCP [23] and UBQLN2 [24]. Two networks were enriched for functional categories related to regulation of apoptosis (**Table 2**). Dysregulation of pathways related to apoptosis has also been previously implicated in ALS [25]. Both protein processing and regulation of apoptosis have been linked to ALS in the context of ER stress [18]; misfolded protein accumulation can induce ER stress, and chronic ER stress can lead to apoptosis.

A network of genes enriched for 'RNA splicing' as was up-regulated in both cell types under examination suggesting that it may represent an upstream effect of the expansion. Analysis of the 'RNA splicing' network signal in both models showed that the similarity extended beyond the functional enrichment to the actual genes dysregulated. Moreover the dysregulated gene lists were significantly enriched with independently generated candidate protein binding partners of the GGGGCC-repeat expansion from our own work [8] and that of others [7,26,27]. We have previously proposed that *C9ORF72*-disease involves dynamic sequestration of a significant number of RBPs involved in mRNA splicing, by RNA foci transcribed from the GGGGCC repeat [8]. The observed up-regulation of genes encoding these proteins in both *C9ORF72*-ALS motor neurons and lymphoblastoid cells is consistent with attempted compensation by the cell for a sequestration process.

Exon splicing in *C9ORF72*-ALS lymphoblastoid cells

In view of these findings we attempted to examine global splicing function within the cell. Despite their being non-neuronal, we utilised lymphoblastoid cells in this analysis because of the large number of samples available and the accessibility of high quality RNA. It is reasonable to expect that a molecular phenotype observed in lymphoblastoid cells might also be present in the central nervous system (CNS). We have previously shown that detectable *C9ORF72* expansion length [28] and transcription of RNA foci [8] are comparable between lymphoblastoid cell lines and the CNS.

Splicing errors are likely to be a normal occurrence for which the cell is able to compensate [13]. However, if the load of these errors is increased then the compensatory mechanism may be overcome and the probability of this occurring might be expected to increase with time. This is consistent with the late age of onset and markedly variable phenotype found in *C9ORF72*-disease. In order to quantify splicing errors, we defined functionally appropriate splicing as likely to be consistent between members of a particular group: *C9ORF72*-ALS, non-*C9ORF72* ALS or controls. We identified a reduction in splicing consistency, or an increase in the splicing error rate, in *C9ORF72*-ALS samples compared to non-*C9ORF72* ALS samples and controls; moreover the splicing error rate was higher in samples derived from *C9ORF72*-ALS patients with shorter survival compared to samples derived from *C9ORF72*-ALS patients with

a longer survival, suggesting a link with CNS toxicity. Consistent with our hypothesis there was a trend for the frequency of sense RNA foci to be higher in lymphoblastoid cell lines derived from patients with a shorter survival; this might be expected to increase the sequestration of RNA splicing proteins and thus exacerbate the production of splicing errors.

A number of studies have previously examined the transcriptome in the presence of expanded *C9ORF72* [3–5,29]; the findings of these studies have so far been inconsistent. An increase in the number of splicing errors is a potential explanation for this finding. We await further validation of our results using new technologies such as RNA sequencing and utilising newly emerging disease models including iPS derived motor neurons from patients with *C9ORF72* mutations.

Materials and Methods

Transcriptome analysis

Laser captured motor neurones. Brain and spinal cord tissue from eight *C9ORF72*-ALS patients and three neurologically normal human control subjects was obtained from the Sheffield Brain Tissue Bank (**Table 3**). *C9ORF72*-ALS samples were identified by repeat-primed PCR of the *C9ORF72* gene [1,2]. Clinically these patients resembled the full clinical spectrum of *C9ORF72*-ALS: Mean age of onset was 61 years (range 56 to 66 years) and mean disease duration was approximately 2 years (range 7 months to 43 months). Tissue donated for research was obtained with written informed consent from the next of kin, and in accordance with the UK Human Tissue Authority guidelines on tissue donation. The work was approved by the South Yorkshire Ethics Committee.

Spinal cord sections from the limb enlargements were collected postmortem, processed according to standard protocols [30], and stored at −80ºC until required. Cervical spinal cord sections were prepared, between 800 and 1200 motor neurons were isolated and RNA was extracted using methods described previously [31]. RNA quantity and quality was assessed on the Nanodrop spectrophotometer and Agilent Bioanalyser, respectively, to ensure all samples were of comparable and sufficient quality to proceed. RNA (20–25ng) was linearly amplified using the Affymetrix Two Cycle cDNA synthesis protocol to produce biotin-labelled copy RNA. Copy RNA (15μg) was fragmented for 15min and hybridized to the Human Genome U133 Plus 2.0 GeneChips, according to Affymetrix protocols. Array washing and staining was

Table 3. Clinical information relating to motor neurons laser captured from ALS patients and controls, utilised in gene level microarray analysis.

Sample Type	Gender	Age	Duration	Diagnosis	Presentation	C9orf72
Control1	F	52	-	-	-	-
Control2	M	63	-	-	-	-
Control3	F	65	-	-	-	-
Patient1	F	62	2.00	Familial	Bulbar	+
Patient2	F	61	3.33	Sporadic	Bulbar	+
Patient3	M	66	1.17	Familial	Bulbar	+
Patient4	F	56	3.58	Familial	Limb	+
Patient5	M	62	1.67	Sporadic	Bulbar	+
Patient6	F	61	3.50	Sporadic	Limb	+
Patient7	M	70	2.17	Familial	Limb	+
Patient8	F	58	0.58	Sporadic	Limb	+

Age at symptom onset and disease duration is provided in years. Abbreviations: M = male, F = female.

doi:10.1371/journal.pone.0127376.t003

performed in the GeneChip fluidics station 400 and arrays were scanned on the GeneChip 3000 scanner. GeneChip Operating Software (GCOS) was used to generate signal intensities for each transcript.

Data Analysis. Data were normalised using the Puma package which quantifies technical variability to improve the estimation of gene expression [32, 33]. The next step was to identify networks of genes with correlated expression which are likely to represent functional groups. To reduce the computational burden and enhance the signal strength in the data, genes were ranked by t-statistic in a disease versus control comparison; the top 10,000 genes were then taken forward. For network detection, genes were further filtered to find the 5000 most connected (as quantified by Pearson correlation coefficient) genes; by definition, networked genes are strongly connected and therefore this should not lead to loss of information [34]. Network detection was performed using the weighted gene coexpression network analysis (WGCNA) package [35]. The correlation between expression of a given network of genes and whether a sample was a *C9ORF72*-case or a control was quantified and a Student's asymptotic p-value calculated; p-values <0.05 were taken to be significant. Differentially expressed networks were examined and an enrichment analysis performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) [36,37]. Enrichment was calculated by functional annotation clustering using the 'high' i.e. specific, Gene Ontology 'biological processes' terms.

Lymphoblastoid cell lines. Lymphoblastoid cell lines derived from Caucasian ALS patients (n = 56) and neurologically normal controls (n = 15), all of Northern European descent, were obtained from the UK Motor Neurone Disease Association (MNDA) DNA Bank (**Table 4**). *C9ORF72*-ALS samples were identified by repeat-primed PCR of the *C9ORF72* gene [1,2]. Clinically these patients resembled the full clinical spectrum of *C9ORF72*-ALS: Mean age of onset was 58 years (range 28 to 75 years) and mean disease duration was approximately 2 years (range 2 months to 83 months). All samples were collected with written informed consent from the donor, and the work was approved by the South Yorkshire Ethics Committee.

Total RNA was extracted from ALS patient and control-derived lymphoblastoid cell lines using QIAGEN's RNeasy Mini Kit following the manufacturer's recommendations. A 75μL LCL suspension, containing approximately $5x10^6$ cells, typically yields between 1.9 and 13.6μg total RNA with a mean concentration of approximately 170ng/μl as assessed the by the NanoDrop 1000 spectrophotometer (Thermo Scientific). The quality of the isolated material was analysed using the 2100 bioanalyzer with an RNA 6000 Nano LabChip Kit (Agilent Technologies, Inc.). Linear amplification of RNA with an input of approximately 300ng of starting material was performed using the Ambion Whole Transcript (WT) Expression Assay (*Applied Biosystems*) and Affymetrix GeneChip WT Terminal Labelling Kit. This procedure generated fragments of biotin-labelled sense-stranded copy DNA (6–10μg) between 40 and 70 nucleotides in length that were hybridized onto Human Exon 1.0ST GeneChip Arrays according to Affymetrix protocols. Array washing, staining and visualisation were performed as described for motor neuron derived RNA.

Data analysis. Network analysis of gene expression in the lymphoblastoid cell lines was identical to that in the motor neurons. However, because of the exon level probing, after Puma normalisation, approximately twice as many transcripts were quantified. This was taken into consideration in the filtering steps and the same proportion of transcripts were analysed at each stage rather than the exact same number i.e. the top 20,000 genes ranked by t-statistic were filtered to the 10,000 most connected for the network analysis.

Exon level data were analysed using the 'finding isoforms using robust multichip analysis' (FIRMA) package [38] which itself is part of the Aroma Affymetrix package [39]. The FIRMA step was then applied to detect alternative splicing; a FIRMA score is calculated for each exon. The score represents the result of fitting a transcript-level model to the observed data and

Table 4. Clinical information relating to lymphoblastoid cell lines derived from ALS patients and controls, utilised in exon level microarray analysis.

Sample Type	Gender	Age	Duration	Diagnosis	Presentation	C9orf72
Control1	F	52	-	-	-	-
Control2	M	69	-	-	-	-
Control3	F	65	-	-	-	-
Control4	F	84	-	-	-	-
Control5	M	56	-	-	-	-
Control6	F	59	-	-	-	-
Control7	M	73	-	-	-	-
Control8	F	67	-	-	-	-
Control9	M	47	-	-	-	-
Control10	M	64	-	-	-	-
Control11	F	41	-	-	-	-
Control12	M	36	-	-	-	-
Control13	M	61	-	-	-	-
Control14	M	54	-	-	-	-
Control15	F	63	-	-	-	-
Patient1	F	69	>4.00	Familial	Limb	+
Patient2	F	61	2.96	Familial	Limb	-
Patient3	F	28	1.10	Familial	Bulbar	+
Patient4	M	44	2.11	Familial	Respiratory	-
Patient5	F	46	Unknown	Familial	Bulbar	-
Patient6	M	69	1.76	Familial	Limb	+
Patient7	M	48	Unknown	Familial	Mixed	-
Patient8	M	57	5.71	Familial	Mixed	-
Patient9	F	57	1.21	Familial	Mixed	+
Patient10	M	63	>5.00	Familial	Limb	+
Patient11	F	62	0.17	Familial	Bulbar	+
Patient12	F	64	6.92	Familial	Limb	+
Patient13	M	59	<1.00	Familial	Unknown	+
Patient14	M	63	1.71	Familial	Mixed	+
Patient15	F	56	4.14	Familial	Limb	+
Patient16	M	47	1.63	Familial	Limb	+
Patient17	F	51	0.97	Familial	Bulbar	+
Patient18	F	61	Unknown	Familial	Bulbar	-
Patient19	M	73	1.88	Sporadic	Respiratory	-
Patient20	M	60	1.15	Sporadic	Bulbar	+
Patient21	M	64	2.36	Sporadic	Bulbar	-
Patient22	F	68	3.31	Sporadic	Bulbar	-
Patient23	M	68	1.56	Sporadic	Limb	+
Patient24	F	72	4.66	Sporadic	Limb	+
Patient25	M	58	1.40	Sporadic	Bulbar	-
Patient26	M	54	2.89	Sporadic	Bulbar	-
Patient27	M	53	3.28	Sporadic	Limb	-
Patient28	F	52	2.25	Sporadic	Limb	+
Patient29	M	72	2.58	Sporadic	Limb	-
Patient30	M	60	1.08	Sporadic	Bulbar	-
Patient31	F	67	1.47	Sporadic	Bulbar	+

(*Continued*)

Table 4. (*Continued*)

Sample Type	Gender	Age	Duration	Diagnosis	Presentation	C9orf72
Patient32	F	37	1.74	Sporadic	Limb	+
Patient33	M	56	2.20	Sporadic	Limb	+
Patient34	M	59	1.84	Sporadic	Limb	-
Patient35	F	70	2.13	Sporadic	Limb	-
Patient36	M	38	2.83	Sporadic	Mixed	-
Patient37	M	45	1.47	Sporadic	Limb	+
Patient38	F	48	~4.00	Sporadic	Bulbar	+
Patient39	F	72	1.87	Sporadic	Bulbar	-
Patient40	M	72	0.52	Sporadic	Limb	+
Patient41	F	75	1.05	Sporadic	Limb	-
Patient42	F	52	2.18	Sporadic	Limb	-
Patient43	F	58	1.33	Sporadic	Mixed	+
Patient44	M	47	1.57	Sporadic	Limb	+
Patient45	F	48	5.95	Sporadic	Limb	+
Patient46	M	64	0.66	Sporadic	Limb	+
Patient47	F	37	4.50	Sporadic	Bulbar	+
Patient48	M	70	1.24	Sporadic	Limb	-
Patient49	F	70	3.04	Sporadic	Limb	-
Patient50	M	61	2.57	Sporadic	Bulbar	-
Patient51	M	62	1.96	Sporadic	Limb	+
Patient52	F	58	<1.00	Sporadic	Bulbar	+
Patient53	M	61	~4.00	Sporadic	Mixed	+
Patient54	M	65	1.40	Sporadic	Limb	+

Age at symptom onset and disease duration is provided in years. Abbreviations: M = male, F = female.

doi:10.1371/journal.pone.0127376.t004

observing the disparity between the model and the exon-level intensity of each individual exon. Thus exons with a different level of expression to their parent transcript i.e. those which are spliced in or out, are identified. Utilising all probes specific for an entire transcript results in a significant improvement on the estimation of exon expression compared to using the relatively small number of probes specific to a given exon in isolation. FIRMA scores were log transformed prior to analysis. Highly negative or positive values of the FIRMA score are indicative of alternative exon skipping or inclusion respectively. The 1st and 99th percentiles of the FIRMA score for all exons in all samples were used to identify exons with the most evidence of alternative splicing, as used previously [40].

Consistency of splicing within a sample group such as patients or controls was evaluated by comparing the number of splicing events which occurred in 1, 2, 3. . .n samples within the group. To allow comparison between groups, comparison was made with the situation in which exons are spliced in or out at random. In each case the random situation was modelled with a Poisson distribution and the observed data were fitted to a negative binomial distribution (**S1 Fig**). θ is a quantification of the overdispersion in the negative binomial distribution with respect to an equivalent Poisson distribution, which is therefore a measure of non-random choice, i.e. consistency, in the splicing observed in each sample group. The variance of a negative binomial distribution is given by $\mu+\mu 2/\theta$ where μ = mean. In contrast the variance of a Poisson distribution is equal to μ. Therefore a higher level of θ corresponds to a variance closer to the Poisson distribution and reduced consistency of splicing.

qPCR based validation of transcriptome changes

Total RNA from lymphoblastoid cells was amplified using High Capacity RNA-to-cDNA kit (Applied Biosystems). Quantitative PCR (QPCR) primers for HNRNPF, FUS, HNRNPH2 and RBM3 transcripts were designed using Eurofins online primer design software (http://www.eurofinsdna.com). QPCR of 75 *C9ORF72*-ALS cases and 35 controls was performed using Brilliant II SYBR Green QPCR Master Mix (Stratagene) on the Stratagene 3000, as described previously [41]. RNA from groups of five *C9ORF72*-ALS or five control samples was pooled. These samples were obtained from the UK MNDA DNA Bank and included those samples utilised in the microarray analysis, as well as additional samples. T-tests were used to determine if the relative differences in transcript expression in lymphoblastoid cells between *C9ORF72*-ALS samples and controls were statistically significant.

Estimation of expansion size in lymphoblastoid cell lines

GGGGCC expansion size was estimated using a Southern hybridisation based protocol as previously described [28] using DNA derived from patients with rapid (<2 years, n = 17) or slowly (>4 years, n = 7) progressive disease.

Quantification of abundance of RNA foci in lymphoblastoid cell lines

A 5' TYE-563-labeled LNA (16-mer fluorescent)-incorporated DNA probe was used against the sense (Exiqon, Inc.; batch number 607323) and the antisense RNA hexanucleotide repeat (Exiqon, Inc.; batch number 610331). Slides were prepared and RNA foci were visualised as described previously [8]. More than fifty lymphoblastoid cells derived from patients with rapid (<2 years, n = 3) or slowly (>4 years, n = 3) progressive disease were imaged.

Supporting Information

S1 Fig. Quantification of splicing events shared between lymphoblastoid cell lines of each sample group. Plots of the number of splicing events (y-axis) which were present in a given number of lymphoblastoid cell lines (x-axis) within a particular sample group. Sample groups from top to bottom are: normal controls, *C9ORF72*-ALS patients, non-*C9ORF72* ALS patients, *C9ORF72*-ALS patients with survival <2 years and *C9ORF72*-ALS patients with survival >4 years. In each plot the left-hand line represents a Poisson fit to the observed data i.e. the random case. The right-hand line is the observed data and the dotted line represents the negative binomial distribution fit to the observed data. In each case the negative binomial provides a relatively good fit to the observed data. θ as shown in Fig 4, is a quantification of the overdispersion in the negative binomial compared to the Poisson fit to the observed data i.e. the degree of consistency in the splicing observed in each sample group.
(TIF)

S1 Table. Genes associated with the 'RNA splicing' network signal and within the GO term 'RNA Processing.' Genes listed are within the GO category 'RNA Processing' and are significantly associated ($p<0.05$) with the 'RNA splicing' network signal in either *C9ORF72*+ lymphoblastoid cell lines or motor neurons. A fold change of >1 equates to up-regulation and a fold change of <1 equates to down-regulation.
(XLSX)

Acknowledgments

Samples used in this research were in part obtained from the UK MND DNA Bank for MND Research, funded by the MND Association and the Wellcome Trust. We are grateful to all of the patients with ALS and their family members who donated biosamples for research.

Author Contributions

Conceived and designed the experiments: JCK JRH GH MR JK PJS. Performed the experiments: JCK JJB PRH MW AH CG. Analyzed the data: JCK JJB PRH JRH MR PJS. Contributed reagents/materials/analysis tools: MR JK PRH PJS. Wrote the paper: JCK JJB PRH MW AH CG JRH GH MR JK PJS.

References

1. DeJesus-Hernandez M, Mackenzie I, Boeve B, Boxer A, Baker M, Rutherford NJ, et al. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. Neuron 2011; 72: 245–256. doi: 10.1016/j.neuron.2011.09.011 PMID: 21944778
2. Renton Alan E, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, et al. A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. Neuron 2011; 72: 257–268. doi: 10.1016/j.neuron.2011.09.010 PMID: 21944779
3. Donnelly CJ, Zhang PW, Pham JT, Heusler AR, Mistry NA, Vidensky S, et al. RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. Neuron 2013; 80: 415–428. doi: 10.1016/j.neuron.2013.10.015 PMID: 24139042
4. Sareen D, O'Rourke JG, Meera P, Muhammad AKMG, Grant S, Simpkinson M, et al. Targeting RNA Foci in iPSC-Derived Motor Neurons from ALS Patients with a C9ORF72 Repeat Expansion. Science Translational Medicine 2013; 5: 208ra149.
5. Lagier-Tourenne C, Baughn M, Rigo F, Sun S, Liu P, Li HR, et al. Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. Proc Natl Acad Sci U S A 2013; 110:E4530–E4539. doi: 10.1073/pnas.1318835110 PMID: 24170860
6. Mizielinska S, Lashley T, Norona FE, Clayton EL, Ridler CE, Fratta P, et al. C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. Acta Neuropathol. 2013; 126: 845–857. doi: 10.1007/s00401-013-1200-z PMID: 24170096
7. Lee YB, Chen HJ, Peres JN, Gomez-Deza J, Attig J, Stalekar M, et al. Hexanucleotide Repeats in ALS/FTD Form Length-Dependent RNA Foci, Sequester RNA Binding Proteins, and Are Neurotoxic. Cell Rep. 2013; 5: 1178–1186. doi: 10.1016/j.celrep.2013.10.049 PMID: 24290757
8. Cooper-Knock J, Walsh MJ, Higginbottom A, Highley JR, Dickman MJ, Edbauer D, et al. Sequestration of multiple RNA Recognition Motif-containing proteins by C9ORF72 repeat expansions. Brain 2014; 137:2040–51. doi: 10.1093/brain/awu120 PMID: 24866055
9. Ash PE, Bieniek KF, Gendron TF, Caulfield T, Lin WL, Dejesus-Hernandez M, et al. Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron 2013; 77: 639–646. doi: 10.1016/j.neuron.2013.02.004 PMID: 23415312
10. Mackenzie IR, Arzberger T, Kremmer E, Troost D, Lorenzl S, Mori K, et al. Dipeptide repeat protein pathology in C9ORF72 mutation cases: clinico-pathological correlations. Acta Neuropathol. 2013; 126: 859–879. doi: 10.1007/s00401-013-1181-y PMID: 24096617
11. Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E, et al. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science 2013; 339: 1335–1338. doi: 10.1126/science.1232927 PMID: 23393093
12. Cooper-Knock J, Shaw PJ, Kirby J The widening spectrum of C9ORF72-related disease; genotype/phenotype correlations and potential modifiers of clinical phenotype. Acta Neuropathol. 2014; 127:333–45. doi: 10.1007/s00401-014-1251-9 PMID: 24493408
13. Pickrell JK, Pai AA, Gilad Y, Pritchard JK Noisy splicing drives mRNA isoform diversity in human cells. PLoS Genet. 2014; 6: e1001236.
14. Parakh S, Spencer DM, Halloran MA, Soo KY, Atkin JD Redox regulation in amyotrophic lateral sclerosis. Oxid Med Cell Longev 2013; 2013: 408681. doi: 10.1155/2013/408681 PMID: 23533690
15. Jawaid A, Salamone AR, Strutt AM, Murthy SB, Wheaton M, McDowell EJ, et al. ALS disease onset may occur later in patients with pre-morbid diabetes mellitus. Eur J Neurol. 2010; 17: 733–739. doi: 10.1111/j.1468-1331.2009.02923.x PMID: 20074230

16. Wang W, Bu B, Xie M, Zhang M, Yu Z, Tao D. Neural cell cycle dysregulation and central nervous system diseases. Prog Neurobiol. 2009; 89: 1–17. doi: 10.1016/j.pneurobio.2009.01.007 PMID: 19619927
17. Cooper-Knock J, Kirby J, Ferraiuolo L, Heath PR, Rattray M, Shaw PJ. Gene expression profiling in human neurodegenerative disease. Nat Rev Neurol. 2012; 8: 518–530. doi: 10.1038/nrneurol.2012.156 PMID: 22890216
18. Matus S, Valenzuela V, Medinas DB, Hetz C ER Dysfunction and Protein Folding Stress in ALS. Int J Cell Biol 2013; 2013: 674751. doi: 10.1155/2013/674751 PMID: 24324498
19. Vijayalakshmi K, Alladi PA, Ghosh S, Prasanna VK, Sagar BC, Nalini A, et al. Evidence of endoplasmic reticular stress in the spinal motor neurons exposed to CSF from sporadic amyotrophic lateral sclerosis patients. Neurobiol Dis. 2011; 41: 695–705. doi: 10.1016/j.nbd.2010.12.005 PMID: 21168498
20. Rohrl C, Eigner K, Winter K, Korbelius M, Obrowsky S, Kratky D, et al. Endoplasmic reticulum stress impairs cholesterol efflux and synthesis in hepatic cells. J Lipid Res. 2014; 55: 94–103. doi: 10.1194/jlr.M043299 PMID: 24179149
21. Raman R, Allen SP, Goodall EF, Kramer S, Ponger LL, Heath PR, et al. Gene expression signatures in motor neuron disease fibroblasts reveal dysregulation of metabolism, hypoxia-response and RNA processing functions. Neuropathol Appl Neurobiol. 2015; 41:201–226 doi: 10.1111/nan.12147 PMID: 24750211
22. Blokhuis AM, Groen EJ, Koppers M, van den Berg LH, Pasterkamp RJ Protein aggregation in amyotrophic lateral sclerosis. Acta Neuropathol. 2013; 125: 777–794. doi: 10.1007/s00401-013-1125-6 PMID: 23673820
23. Johnson JO, Mandrioli J, Benatar M, Abramzon Y, Van Deerlin VM, Trojanowski JQ, et al.Exome sequencing reveals VCP mutations as a cause of familial ALS. Neuron 2010; 68: 857–864. doi: 10.1016/j.neuron.2010.11.036 PMID: 21145000
24. Deng HX, Chen W, Hong ST, Boycott KM, Gorrie GH, Siddique N, et al. Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. Nature 2011; 477: 211–215. doi: 10.1038/nature10353 PMID: 21857683
25. Sathasivam S, Shaw PJ Apoptosis in amyotrophic lateral sclerosis—what is the evidence? Lancet Neurol. 2005; 4: 500–509. PMID: 16033692
26. Mori K, Lammich S, Mackenzie IR, Forne I, Zilow S, Kretzschmar H, et al. hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with C9orf72 mutations. Acta Neuropathol. 2013; 125: 413–423. doi: 10.1007/s00401-013-1088-7 PMID: 23381195
27. Haeusler AR, Donnelly CJ, Periz G, Simko EA, Shaw PG, Kim MS, et al. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. Nature 2014; 507: 195–200. doi: 10.1038/nature13124 PMID: 24598541
28. Buchman VL, Cooper-Knock J, Connor-Robson N, Higginbottom A, Kirby J, Razinskaya OD, et al. Simultaneous and independent detection of C9ORF72 alleles with low and high number of GGGGCC repeats using an optimised protocol of Southern blot hybridisation. Mol Neurodegener. 2013; 8: 12. doi: 10.1186/1750-1326-8-12 PMID: 23566336
29. Ismail A, Cooper-Knock J, Highley JR, Milano A, Kirby J, Goodall E, et al. Concurrence of multiple sclerosis and amyotrophic lateral sclerosis in patients with hexanucleotide repeat expansions of C9ORF72. J Neurol Neurosurg Psychiatry 2013; 84: 79–87. doi: 10.1136/jnnp-2012-303326 PMID: 23085936
30. Ince PG, McArthur FK, Bjertness E, Torvik A, Candy JM, Edwardson JA. Neuropathological diagnoses in elderly patients in Oslo: Alzheimer's disease, Lewy body disease, vascular lesions. Dementia 1995; 6: 162–168. PMID: 7620529
31. Ferraiuolo L, Heath PR, Holden H, Kasher P, Kirby J, Shaw PJ. Microarray analysis of the cellular pathways involved in the adaptation to and progression of motor neuron injury in the SOD1 G93A mouse model of familial ALS. J Neurosci. 2007; 27: 9201–9219. PMID: 17715356
32. Pearson RD, Liu X, Sanguinetti G, Milo M, Lawrence ND, Rattray M. puma: a Bioconductor package for propagating uncertainty in microarray analysis. BMC Bioinformatics 2009; 10: 211. doi: 10.1186/1471-2105-10-211 PMID: 19589155
33. Rattray M, Liu X, Sanguinetti G, Milo M, Lawrence ND Propagating uncertainty in microarray data analysis. Brief Bioinform. 2006; 7: 37–47. PMID: 16761363
34. Ghazalpour A, Doss S, Zhang B, Wang S, Plaisier C, Castellanos R, et al. Integrating genetic and network analysis to characterize genes related to mouse weight. PLoS Genet. 2006; 2: e130. PMID: 16934000
35. Langfelder P, Horvath S WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008; 9: 559. doi: 10.1186/1471-2105-9-559 PMID: 19114008

36. Huang DW, Sherman BT, Lempicki RA Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protocols 2008; 4: 44–57.

37. Huang DW, Sherman BT, Lempicki RA Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research 2009; 37: 1–13. doi: 10.1093/nar/gkn923 PMID: 19033363

38. Purdom E, Simpson KM, Robinson MD, Conboy JG, Lapuk AV, Speed TP. FIRMA: a method for detection of alternative splicing from exon array data. Bioinformatics 2008; 24: 1707–1714. doi: 10.1093/bioinformatics/btn284 PMID: 18573797

39. Bengtsson H, Simpson K, Bullard J, Hansen K. aroma.affymetrix: A generic framework in R for analyzing small to very large Affymetrix data sets in bounded memory. 2008; Tech Report #745, Department of Statistics, University of California, Berkeley.

40. Sveen A, Agesen TH, Nesbakken A, Rognum TO, Lothe RA, Skotheim RI. Transcriptome instability in colorectal cancer identified by exon microarray analyses: Associations with splicing factor expression levels and patient survival. Genome Med. 2011; 3: 32. doi: 10.1186/gm248 PMID: 21619627

41. Kirby J, Halligan E, Baptista MJ, Allen S, Heath PR, Holden H, et al. Mutant SOD1 alters the motor neuronal transcriptome: implications for familial ALS. Brain 2005; 128: 1686–1706. PMID: 15872021

Clinical/Scientific Notes

Johnathan Cooper-Knock, BA
Adrian Higginbottom, PhD
Natalie Connor-Robson, MSc
Nadhim Bayatti, PhD
Joanna J. Bury, MSc
Janine Kirby, PhD
Natalia Ninkina, MD, PhD
Vladimir L. Buchman, MD, PhD
Pamela J. Shaw, MD

C9ORF72 TRANSCRIPTION IN A FRONTOTEMPORAL DEMENTIA CASE WITH TWO EXPANDED ALLELES

Discovery of intronic hexanucleotide repeat expansions of the *C9ORF72* gene in a significant proportion of patients with amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD)[1,2] was an important step for research into these disorders. The *C9ORF72* genetic variant is more common than other described mutations, and, unlike patients with mutations in *SOD1*, *C9ORF72*-ALS clinically and pathologically resembles the more numerous sporadic form.[3] However, progress has been limited by lack of understanding of the function of the *C9ORF72* locus in health and disease. It is unknown whether the expansion causes disease by a gain of toxicity, whether it disrupts expression of the wild-type protein encoded by the *C9ORF72* gene, or some combination of both mechanisms.[1,2,4]

Case. Our case is a woman who presented with deteriorating handwriting at age 58 years. Later she developed features of frontal dysfunction and parkinsonism; she received a formal diagnosis of behavioral-variant FTD from a consultant neurologist. Two years after diagnosis, she has not developed motor weakness or denervation changes on EMG. The patient had a brother who died of ALS at age 63 years; no other family members are known to have had neurologic disease, although available information is limited. Several relatives died relatively young (<50 years) from non-neurologic causes, including her parents; she has no children.

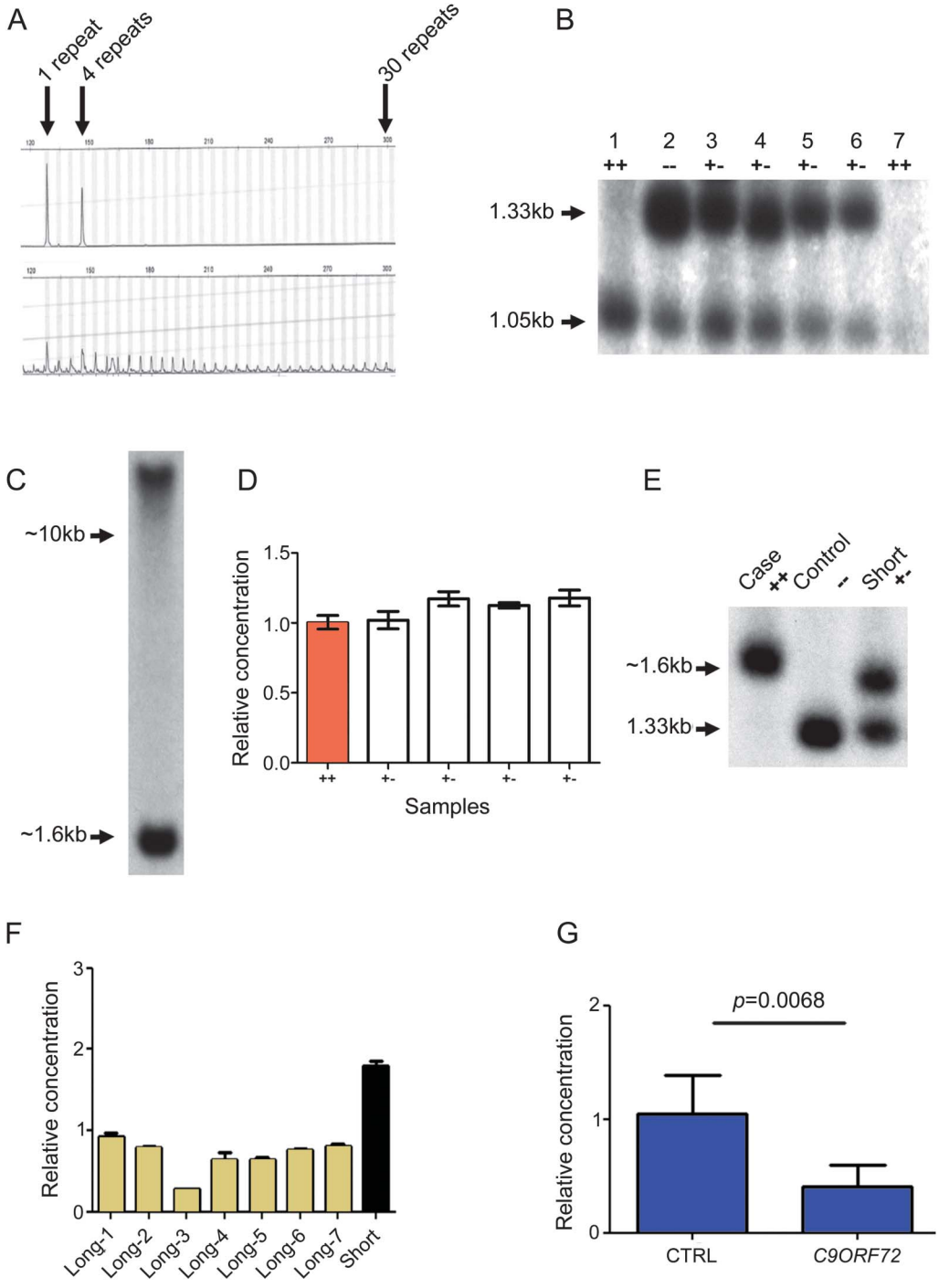
***Detection of 2 expanded C9ORF72* alleles.** PCR analysis of genomic DNA extracted from our patient's venous blood did not detect a normal-length *C9ORF72* allele (figure, A); similarly, Southern hybridization analysis[5] revealed no normal-length *C9ORF72* allele in venous blood or saliva (figure, B). Instead, 2 expanded alleles were detected of 50 ± 5 repeats and >2,000 repeats (figure, C).

***Investigation of C9ORF72* mRNA expression.** qRT-PCR analysis of RNA extracted from venous blood cells in our patient and several patients with expansions of >2,000 repeats in 1 *C9ORF72* allele and the absence of expansion (i.e., <20 repeats) in another allele revealed similar levels of *C9ORF72* mRNA in all samples (figure, D). Southern hybridization identified a lymphoblastoid cell line derived from a *C9ORF72*-ALS patient with one expanded *C9ORF72* allele of similar size (45 ± 5 repeats) to the smaller expansion in our case (figure, E). qRT-PCR of RNA extracted from this cell line and other lines with expansions of >2,000 repeats in 1 *C9ORF72* allele revealed that *C9ORF72* mRNA levels were twice as high in the line with the smaller expansion (figure, F). Ethics committee approval and written consent was obtained for all biosamples.

Discussion and conclusions. Identification of a patient with 2 expanded *C9ORF72* alleles is an important step in the study of *C9ORF72* disease. Her disease severity, defined by age at onset and disease duration, is not remarkably different from other *C9ORF72*-positive patients.[6] Notably, the length of expansion is different in her 2 *C9ORF72* alleles, which allowed us to explore the effect of shorter expansions on *C9ORF72* mRNA expression. We and others have shown that the majority of *C9ORF72* neurodegeneration patients carry a repeat expansion of >2,000 repeats,[5] although it has been suggested that more than 30[1,2] or even as little as 20 to 22 repeats in 1 *C9ORF72* allele are pathogenic.[7] A recent study has reported hypermethylation of a CpG island 5′ to the repeat sequence, which did not occur in samples with intermediate-length expansions of up to 43 repeats.[8] If this is the mechanism underlying reduced mRNA expression, then it should not affect smaller repeat sizes.

Previously published[2] and our own data (figure, G) demonstrate that the level of *C9ORF72* mRNA is reduced in blood cells from patients with long expansions of >2,000 repeats, suggesting that expression of RNA from the expanded allele is compromised. In a lymphoblastoid cell line from a patient with a short expansion in 1 *C9ORF72* allele, estimated as 45 ± 5 repeats by Southern hybridization, the level of *C9ORF72* mRNA was approximately double the level in cell lines from patients with long expansions (figure, F). This suggests that an allele with a short pathogenic expansion is normally expressed, but this finding might reflect a compensatory increase in transcription of the normal allele. However, since our patient with 2 expansions does not possess a normal allele, such

Figure ***C9ORF72*** **expansion size and** ***C9ORF72*** **mRNA expression in lymphoblastoid cells and venous blood**



(A) Genotyping PCR of a wild-type control and our patient. The shaded lines represent numbers of repeats from 1 to 30. Thus the upper panel shows a heterozygous control with 2 normal-length alleles of 1 and 4 repeats. No normal-length allele of less than 30 repeats is detected in our patient, as shown in the lower panel. (B, C) Southern hybridization–based detection of the *C9ORF72* allele. (B) Analysis of DNA extracted from venous blood of 5 *C9ORF72*-positive patients and 1 control. The 1.33-kb band corresponds to an *Eco*RI/*Xba*I fragment derived from a nonexpanded locus. This band is present in patients with a single *C9ORF72* expansion (lanes 3–6, +−) and in normal controls without an expansion (lane 2, −−) but is absent in the patient with 2 expanded alleles (lanes 1 and 7, ++) in both venous blood (lane 1) and saliva (lane 7). The 1.05-kb band is an internal control to show that the absence of 1.33-kb band is not due to a low amount of DNA loaded or its inability to hybridize with the labeled probe. (C) A longer gel allows sizing of both the alleles in the patient with 2 expanded alleles in venous blood. Bands are seen at ~1.6 kb and >12 kb, suggesting that 1 *C9ORF72* allele of the patient carried 50 ± 5 repeats and the other >2,000 repeats. (D) qRT-PCR for *C9ORF72* mRNA in venous blood cells from *C9ORF72*-positive patients; error bars illustrate 95% confidence intervals. Concentration is plotted relative to the concentration in the case with 2 expanded alleles (++). (E) Southern hybridization–based detection of small-size expansion in the *C9ORF72* allele. DNA was extracted from venous blood derived from our case with 2 expanded alleles (left lane), a normal control (middle lane), and a patient with expansion in 1 *C9ORF72* allele of similar length to the

compensation could not explain why *C9ORF72* mRNA expression in her blood cells is equivalent to patients carrying 1 normal allele (figure, D). Thus we conclude that the presence of ~50 copies of the repeat does not significantly affect *C9ORF72* gene transcription or mRNA stability in vivo. If shorter repeats are indeed pathogenic, our evidence suggests that this is unlikely to be mediated by haploinsufficiency.

From the Sheffield Institute for Translational Neuroscience (SITraN) (J.C.-K., A.H., N.B., J.J.B., J.K., P.J.S.), University of Sheffield; and the School of Biosciences (N.C.-R., N.N., V.L.B.), Cardiff University, UK.

Author contributions: The study was conceived and designed by J.C.-K., V.B., and P.J.S. Data acquisition was carried out by J.C.-K., A.H., N.C.-R., N.B., J.J.B., N.N., and V.L.B. Data analysis and interpretation was performed by J.C.-K., A.H., N.C.-R., N.B., J.J.B., J.K., N.N., V.L.B., and P.J.S. The manuscript was critically revised by J.C.-K., A.H., V.L.B., and P.J.S. The study was supervised by N.N., V.L.B., and P.J.S.

Study funding: Supported by grants from the Wellcome Trust (075615/Z/04/z) to V.L.B. and EU Framework 7 (Euromotor No259867) and the SOPHIA project (funded by JPND and MRC) to P.J.S. and J.K. J.C.-K. holds an MND Association/MRC Lady Edith Wolfson Fellowship award (MR/K003771/1). P.J.S. is an NIHR Senior Investigator. Biosample collection was supported by the MND Association and the Wellcome Trust (P.J.S.).

Disclosure: J. Cooper-Knock is supported by an MND Association/MRC Lady Edith Wolfson fellowship award (MR/K003771/1). J. Kirby is supported by an FP7 grant EuroMOTOR (no. 259867) and by Joint Programme for Neurodegenerative Disease (JPND) grant SOPHIA. P. Shaw is supported by NIHR as a Senior Investigator and an FP7 grant EuroMOTOR (no. 259867) and Joint Programme for Neurodegenerative Disease (JPND) grant SOPHIA. V. Buchman is supported by a research grant from the Wellcome Trust (075615/Z/04/z). A. Higginbottom, N. Connor-Robson, N. Bayatti, J. Bury, and N. Ninkina report no disclosures. Go to Neurology.org for full disclosures.

Received March 31, 2013. Accepted in final form July 2, 2013.
Correspondence to Dr. Shaw: pamela.shaw@sheffield.ac.uk

1. Renton AE, Majounie E, Waite A, et al. A hexanucleotide repeat expansion in *C9ORF72* is the cause of chromosome 9p21-Linked ALS-FTD. Neuron 2011;72:257–268.
2. DeJesus-Hernandez M, Mackenzie I, Boeve B, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of *C9ORF72* causes chromosome 9p-linked FTD and ALS. Neuron 2011;72:245–256.
3. Cooper-Knock J, Hewitt C, Highley JR, et al. Clinico-pathological features in amyotrophic lateral sclerosis with expansions in *C9ORF72*. Brain 2012;135:751–764.
4. Polymenidou M, Lagier-Tourenne C, Hutt KR, Bennett CF, Cleveland DW, Yeo GW. Misregulated RNA processing in amyotrophic lateral sclerosis. Brain Res 2012;1462: 3–15.
5. Buchman VL, Cooper-Knock J, Connor-Robson N, et al. Simultaneous and independent detection of C9ORF72 alleles with low and high number of GGGGCC repeats using an optimised protocol of Southern blot hybridisation. Mol Neurodegener 2013;8:12.
6. Majounie E, Renton AE, Mok K, et al. Frequency of the *C9ORF72* hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. Lancet Neurol 2012; 11:323–330.
7. Gomez-Tortosa E, Gallego J, Guerrero-Lopez R, et al. *C9ORF72* hexanucleotide expansions of 20-22 repeats are associated with frontotemporal deterioration. Neurology 2013;80:366–370.
8. Xi Z, Zinman L, Moreno D, et al. Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. Am J Hum Genet Epub 2013 May 22.

smaller expansion in our case (right lane). Bands correspond to genomic fragments carrying 50 ± 5 repeats (left lane), 45 ± 5 repeats (right lane, top band), or no repeats (middle lane and right lane, bottom band). (F) qRT-PCR for *C9ORF72* mRNA in lymphoblastoid cells derived from patients carrying 1 *C9ORF72* allele with repeat expansion. Levels of *C9ORF72* mRNA are approximately double in the patient with a shorter repeat length compared to patients with >2,000 repeats; error bars illustrate 95% confidence intervals. (G) qRT-PCR for *C9ORF72* mRNA in venous blood cells from *C9ORF72*-positive patients and age- and sex-matched controls (CTRL). Levels of *C9ORF72* mRNA are significantly reduced in the patients compared to controls.

4. *C9orf72*-disease modifiers

In section 3 we demonstrated that *C9orf72*-disease encompasses the full spectrum of sporadic disease. This suggests that a number of disease modifiers are present. Each modifier has the potential to be a therapeutic target. The recent observation of a set of monozygotic twins with identical *C9orf72* expansions but different clinical presentations (Xi *et al.* 2014) suggests the environment has a role to play, but work thus far has focused on genetic modifiers as explained below.

4.1. Expansion length

Two of the three prominent proposed pathogenic mechanisms discussed in section 4 are gain-of-function toxicities: RNA toxicity mediated via RNA foci formed from the transcribed repeat sequence, and protein toxicity mediated by DPRs translated from the repeat-RNA. In a gain-of-function scenario, it would be predicted that disease severity is proportional to repeat length. However, a conclusive relationship between repeat length and disease severity has not yet been demonstrated.

4.1.1 Development of Southern blotting protocol

Measurement of repeat length initially proved technically challenging as the large GC rich region is not amenable to standard PCR based sequencing. However, several groups, including ourselves as reported in **paper 9**, have now optimised Southern hybridisation based techniques (Beck *et al.* 2013, Buchman *et al.* 2013, van Blitterswijk *et al.* 2013).

Our probe, reported in **paper 9** (Buchman *et al.* 2013), is targeted to the region of the GGGGCC-repeat expansion including an internal *EcoRI* site. Upon digestion with *EcoRI* and *XbaI* the probe is split into two fragments – the larger portion hybridises to a genomic fragment including the repeat expansion and the smaller

portion hybridises to a genomic fragment within the *C9orf72* intron 1 but downstream from the expansion. When Southern blotting is performed this probe provides an internal standard band for each sample, the size and intensity of which is independent of heterogeneity in the repeat expansion. This is useful since heterogeneity is common and can limit detection of the expanded allele (Buchman *et al.* 2013). Furthermore in **paper 5** we demonstrated significant somatic heterogeneity of expansion length between various tissues derived from the same *C9orf72*-ALS patient(s).

4.1.2 Effect of the expansion length on disease phenotype

Currently, in a pure ALS group, no aspect of the disease phenotype has been shown to significantly correlate with the length of the expansion, regardless of the tissue tested (Dols-Icardo *et al.* 2013, van Blitterswijk *et al.* 2013). In FTD, a direct correlation between repeat size in the frontal cortex and age of onset has been demonstrated, and in the cerebellum a threshold repeat size has been associated with reduced survival (van Blitterswijk *et al.* 2013). This report also indicated that the repeat length in the cerebellum was shorter than in other CNS areas. It is hypothesised that repeat expansions can increase in size through a human lifetime resulting in significant somatic heterogeneity (Clark *et al.* 2007, Buchman *et al.* 2013) and therefore, perhaps the minimum repeat length in the CNS is more reflective of the germline repeat number. If so, the expansion length in the cerebellum may best represent the repeat length which initiated the disease pathogenesis, and the correlation with age of onset in frontal cortex may simply reflect the patient's age.

An alternative way to approach this problem is to look for evidence of anticipation in *C9orf72*-related disease which would be highly suggestive of a relationship between repeat size and disease severity. A series of *C9orf72*-FTD (Benussi *et al.* 2013) and

C9orf72-ALS families (Chio *et al.* 2012) have been described with 7-10 years of anticipation between generations.

Against a direct relationship between repeat length and disease severity is the description of patients with significant clinical disease but a relatively small repeat lengths of <30 units (Byrne *et al.* 2013, Gomez-Tortosa *et al.* 2013). Repeat lengths of 7-24 units are associated with the 9p21 risk haplotype (van der Zee *et al.* 2013). In **paper 10** we describe a patient with clinical ALS, an intermediate expansion of sixteen GGGGCC repeats and the 9p21 risk haplotype, but without the typical neuropathology associated with *C9orf72* disease including RNA foci, DPR inclusions and TDP-43-negative, p62 positive neuronal inclusions in extra-motor areas (Beer *et al.* 2014). As previously stated, we suggest that the pathological characterisation of this case indicates that this individual actually suffered from non-*C9orf72* ALS and the intermediate length expansion was not sufficient to initiate typical *C9orf72*-mediated neuronal injury. This would potentially explain the discordance between gain-of-function models and patients with relatively small expansions. Consistent with this, the frequency of 7-24 repeats, unlike that of longer repeats, is equivalent in patients and controls (van der Zee *et al.* 2013). That patients and controls with intermediate length expansions also tend to carry the 9p21 risk haplotype may reflect the suggestion made earlier, that the risk haplotype predisposes the region to expand; but we suggest that a minimum length, probably greater than 30 repeats, is necessary to initiate *C9orf72*-neuropathology. In this context, it is also interesting that in **paper 8** we demonstrated that expansions of ~50 repeats, unlike longer expansions, are not sufficient to significantly reduce levels of *C9orf72* mRNA (Cooper-Knock *et al.* 2013).

4.2. TMEM106B

A genome-wide association study identified single nucleotide polymorphisms (SNPs) in TMEM106B as a risk factor for FTD with TDP-43 positive pathology (including *C9orf72*-FTD) (Van Deerlin *et al.* 2010). The protein product of TMEM106B is localised to the lysosome. In **paper 11** the haplotype associated with higher risk of FTD-TDP, more particularly the major, or T, allele of rs1990622, was investigated in the context of *C9orf72*-disease (Gallagher *et al.* 2013). The major allele is present at a higher than control frequency in patients with *C9orf72*-FTD and is associated with an earlier age of onset in *GRN*-related FTD. However, in patients with *C9orf72*-FTD, the major allele is associated with a later age of onset and death.

This fascinating complexity might be consistent with similarity in the functions of the *C9orf72* and TMEM106B proteins. Notably membrane trafficking, the previously stated possible role of C9orf72 as a DENN-protein, is a component of lysosome function. It has been suggested that the protective isoform of TMEM106B is expressed at a lower level because of increased degradation mediated via altered glycosylation (Nicholson *et al.* 2013). Interestingly, in contrast to the *C9orf72*-FTD findings; it has been shown that neither TMEM106B allele is significantly associated with a *C9orf72*-related ALS presentation (van Blitterswijk *et al.* 2014). Why the TMEM106B genotype modifies the risk of one phenotype and not the other is unknown, but this suggests that the mechanism of neurotoxicity may be different in each case.

4.3. Other identified modifiers

Another study aiming to identify genetic modifiers of *C9orf72*-disease studied genetic risk factors already associated with ALS in *C9orf72*-expansion carriers and controls. These included altered copy number of *SMN1* and *SMN2*, CAG-repeat expansion of *ATXN2* and GCG-repeat expansion of *NIPA1* (van Blitterswijk *et al.* 2014). Only

ATXN2 expansions of >27 units were present at a higher rate in the *C9orf72* expansion carriers, as in ALS more generally (Elden *et al.* 2010). When this result was broken down by phenotype, it was striking that intermediate length ATXN2 expansions were present in 2-3% of ALS or ALS/FTD patients but were absent in FTD patients, suggesting that expansion of *ATXN2* may predispose *C9orf72*-expansion carriers to develop ALS or ALS/FTD rather than pure FTD. A recent study has confirmed these findings in a larger cohort of ALS, ALS/FTD and FTD patients (Lattante *et al.* 2014). It has been suggested that polyglutamine expansion of *ATXN2* increases the stability of the protein, enhances its interaction with TDP-43, and may promote cytoplasmic mislocalisation of TDP-43 (Elden *et al.* 2010). TDP-43 pathology is a feature of both ALS and FTD and it remains to be discovered why expansions of ATXN2 predispose to ALS and not FTD. The fact that this effect appears to be present independently of *C9orf72* expansion suggests that ATXN2 may have an impact on the 'final common pathway' of disease.

List of Papers:

Paper 8:

Cooper-Knock, J., A. Higginbottom, N. Connor-Robson, N. Bayatti, J. J. Bury, J. Kirby*, et al.* (2013). "*C9ORF72* transcription in a frontotemporal dementia case with two expanded alleles." Neurology

- First author

- Collected samples and performed Southern blotting and data analysis

- Drafted all sections of manuscript

Paper 9:

Buchman V. L., **Cooper-Knock J.**, Connor-Robson N., Higginbottom A., Kirby J., Razinskaya O.D., et al. (2013). "Simultaneous and independent detection of *C9ORF72* alleles with low and high number of GGGGCC repeats using an optimised protocol of Southern blot hybridisation." Mol Neurodegener 8: 12.

- Collected samples and performed Southern blotting

- Drafted introduction, results and dsicussion sections of manuscript

Paper 10:

Beer AM, **Cooper-Knock J**, Higginbottom A, Highley JR, Wharton SB, Ince PG, Milano A, Jones AA, Al-Chalabi A, Kirby J, Shaw PJ. Intermediate length C9orf72 expansion in an ALS patient without classical C9orf72 neuropathology. **Amyotroph Lateral Scler Frontotemporal Degener. [in press]**

- Performed Southern blotting

- Edited manuscript

Paper 11:

Gallagher MD, Suh E, Grossman M, Elman L, McCluskey L, Van Swieten JC, Al-Sarraj S, Neumann M, Gelpi E, Ghetti B, Rohrer JD, Halliday G, Van Broeckhoven C, Seilhean D, Shaw PJ, Frosch MP, Alafuzoff I, Antonell A, Bogdanovic N, Brooks W, Cairns NJ, **Cooper-Knock J**, Cotman C, Cras P, Cruts M, De Deyn PP, Decarli

C, Dobson-Stone C, Engelborghs S, Fox N, Galasko D, Gearing M, Gijselinck I, Grafman J, Hartikainen P, Hatanpaa KJ, Highley JR, Hodges J, Hulette C, Ince PG, Jin LW, Kirby J, Kofler J, Kril J, Kwok JB, Levey A, Lieberman A, Llado A, Martin JJ, Masliah E, McDermott CJ, McKee A, McLean C, Mead S, Miller CA, Miller J, Munoz DG, Murrell J, Paulson H, Piguet O, Rossor M, Sanchez-Valle R, Sano M, Schneider J, Silbert LC, Spina S, van der Zee J, Van Langenhove T, Warren J, Wharton SB, White Iii CL, Woltjer RL, Trojanowski JQ, Lee VM, Van Deerlin V, Chen-Plotkin AS. TMEM106B is a genetic modifier of frontotemporal lobar degeneration with C9orf72 hexanucleotide repeat expansions. **Acta Neuropathol**. 2014;127(3):407-18.

- Responsible for collecting samples and clinical data from Sheffield cohort

MOLECULAR NEURODEGENERATION

METHODOLOGY Open Access

Simultaneous and independent detection of C9ORF72 alleles with low and high number of GGGGCC repeats using an optimised protocol of Southern blot hybridisation

Vladimir L Buchman[1,4*], Johnathan Cooper-Knock[2], Natalie Connor-Robson[1], Adrian Higginbottom[2], Janine Kirby[2], Olga D Razinskaya[3], Natalia Ninkina[1,4] and Pamela J Shaw[2]

Abstract

Background: Sizing of GGGGCC hexanucleotide repeat expansions within the *C9ORF72* locus, which account for approximately 10% of all amyotrophic lateral sclerosis (ALS) cases, is urgently required to answer fundamental questions about mechanisms of pathogenesis in this important genetic variant. Currently employed PCR protocols are limited to discrimination between the presence and absence of a modified allele with more than 30 copies of the repeat, while Southern hybridisation-based methods are confounded by the somatic heterogeneity commonly present in blood samples, which might cause false-negative or ambiguous results.

Results: We describe an optimised Southern hybridisation-based protocol that allows confident detection of the presence of a *C9ORF72* repeat expansion alongside independent assessment of its heterogeneity and the number of repeat units. The protocol can be used with either a radiolabeled or non-radiolabeled probe. Using this method we have successfully sized the *C9ORF72* repeat expansion in lymphoblastoid cells, peripheral blood, and post-mortem central nervous system (CNS) tissue from ALS patients. It was also possible to confidently demonstrate the presence of repeat expansion, although of different magnitude, in both *C9ORF72* alleles of the genome of one patient.

Conclusions: The suggested protocol has sufficient advantages to warrant adoption as a standard for Southern blot hybridisation analysis of GGGGCC repeat expansions in the C9ORF72 locus.

Keywords: C9ORF72, Amyotrophic lateral sclerosis, Southern hybridisation

Background

Hexanucleotide repeat expansion in the *C9ORF72* locus has been identified as a genetic cause, or at least a strong risk factor, for a significant proportion of amyotrophic lateral sclerosis cases [1,2]. It is unknown whether the expansion causes neuronal injury through a toxic gain of function, haploinsufficiency or both mechanisms. Recent studies suggested that not a protein encoded by the *ORF72* gene but dipeptide products of expanded repeat region translation might be toxic for neurons [3,4]. Gain of function is consistent with apparently autosomal dominant inheritance, parallels with other neurodegenerative disorders caused by an intronic expansion [5], and the suggestion of anticipation [6]. Difficulty in estimating the size of the *C9ORF72* expansion has precluded investigation of possible correlations between the repeat length and disease characteristics such as age of onset, severity, or speed of progression.

A repeat primed PCR technique quickly and reliably determines whether a pathological *C9ORF72* expansion of >30 repeats is present in a DNA sample, [1,2] but does not allow even approximate quantification of the repeat number because the 100% GC content of the repeat

* Correspondence: buchmanvl@cf.ac.uk
[1]School of Biosciences, Cardiff University, Museum Avenue, Cardiff CF10 3AX, UK
[4]Institute of Physiologically Active Compounds of RAS, 1 Severniy Proezd, Chernogolovka 142432, Moscow Region, Russian Federation
Full list of author information is available at the end of the article

sequence precludes PCR through the region. Conversely, commonly used Southern hybridisation protocols for detecting fragments encompassing the region of expansion in the *C9ORF72* locus often produce false-negative results when DNA extracted from peripheral blood is analysed. This is due to high repeat number heterogeneity in these samples, which leads to the appearance of multiple high molecular mass fragments forming a smear that might be difficult to distinguish from non-specific binding of the hybridisation probe to digested genomic DNA, even when stringent hybridisation/washing conditions are employed [7-9]. Here we describe a reproducible protocol for unambiguous detection and sizing of the *C9ORF72* repeat expansion by Southern hybridisation.

Results and discussion

Our protocol produces an internal standard band on Southern blots, the size and intensity of which is independent of heterogeneity in the repeat expansion. This is achieved by using as a hybridisation probe a cloned genomic fragment encompassing an internal EcoRI site located close to the repeat expansion region and digesting genomic DNA with two enzymes, EcoRI and XbaI (Figure 1A). The EcoRI site splits the probe into two unequal fragments. Products of a labeling reaction originating from the shorter fragment of the probe hybridise with a 1.05 kb XbaI-EcoRI genomic fragment that does not include the repeat region, while those originating from the longer fragment (~2/3 of the probe length) hybridise with a 1.33 kb XbaI-EcoRI fragment derived from a non-expanded locus, or a larger fragment from a locus with the repeat expansion (Figure 1A).

Subjects without a pathological C9ORF72 hexanucleotide expansion

Because the target XbaI-EcoRI fragments are quite similar in size (thus, no effect of the transfer efficiency during capillary blotting) but the ratio of probe length for detecting each fragment is approximately 1:2, the ratio of intensity of the two bands in the absence of an expanded allele, should be approximately 1:2. Indeed DNA samples from peripheral blood (Figure 1B, lanes 2, 4, 6) or lymphoblastoid cell lines (Figure 1C, lanes 2, 23 and Additional file 1: Table S1) from ALS patients without the repeat expansion show this pattern.

A swing of the intensity ratio of the bands to approximately 1:1 is an indicator of heterozygosity in the *C9ORF72* locus. If the difference in the number of repeats in the two alleles is small (i.e. no pathological repeat expansion in either allele) two close and equally weighted bands are present on Southern blots even when whole blood cell DNA is analysed, as illustrated in Figure 1B lane 7, using a

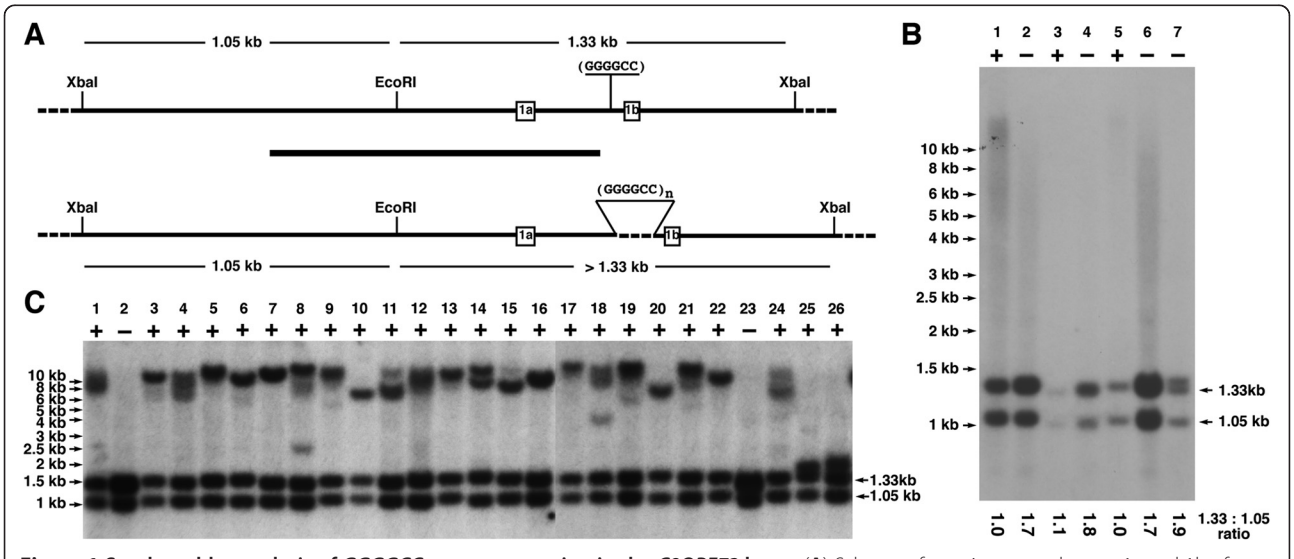


Figure 1 Southern blot analysis of GGGGCC repeat expansion in the *C9ORF72* locus. (**A**) Scheme of a region around exons 1a and 1b of human *C9ORF72* gene with a single copy of GGGGCC sequence (top) or expansion of this repeat (bottom). The black bar in the middle denotes a DNA probe used for Southern blot hybridisation. The sizes of fragments produced by double EcoRI and XbaI digestion and detected by hybridisation with this probe are shown. (**B**) Hybridisation of EcoRI and XbaI digested DNA extracted from the whole blood of patients positive (+) or negative (–) for GGGGCC repeat expansion according to results of the repeat-PCR analysis. The ratios of band intensities obtained by scanning the X-ray film are shown below the image. Note that for sample 7 the intensity used for calculation is combined intensities of two close ~1.33 kb size bands. (**C**) Hybridisation of EcoRI and XbaI digested DNA extracted from the cultured lymphoblastoid cell lines of patients positive (+) or negative (–) for GGGGCC repeat expansion according to results of the repeat-PCR analysis.

DNA sample from a patient carrying one allele with a single copy of the GGGGCC sequence and another allele with 15 copies.

Subjects with a pathological C9ORF72 hexanucleotide expansion in one allele

In cases with a pathological repeat expansion in one allele, a band of larger size is detected in DNA samples extracted from patients' lymphoblastoid cell lines. Examples of these can be found in Figure 1C lanes 10 and 20 (~6 kb band corresponds to an allele with ~750 repeats), 15 (~7.5 kb band corresponds to an allele with ~1000 repeats), 16 and 22 (~8.5 kb band corresponds to an allele with ~1200 repeats), 6 (~10 kb band corresponds to an allele with ~1500 repeats), 3, 7 and 13 (~12 kb band corresponds to an allele with ~1800 repeats), 5, 8, 17 (>12 kb band corresponds to an allele with >2000 repeats). For accuracy of size estimates see Methods section. The presence in the sample of more than one of these larger bands (for example, lanes 4, 8, 14, 18 in Figure 1C, lanes 4, 5 in Figure 2) might indicate instability of the GGGGCC repeat region in the cultured lymphoblastoid cells. Alternatively it might reflect instability and somatic heterogeneity of the repeat region in ALS patients' lymphoid cells, coupled with non-monoclonal origin of the analysed lymphoblastoid cell lines. In whole blood DNA such somatic heterogeneity is clearly evident in some samples, leading to the appearance of multiple bands that coalesce into an ambiguous smear on Southern blots (Figure 1B, lanes 1, 5). For two other samples in the same panel (Figure 1B, lanes 2, 6) lower quality of DNA caused appearance of similar smears but the ratio of 1.05 kb and 1.33 kb band intensities in these cases are close to 1:2 suggesting that both alleles of the C9ORF72 locus have small number of GGGGCC repeats. Therefore, even in those cases with an ambiguous smear, the presence of a pathological allele could be immediately predicted from the ~1:1 ratio of 1.05 kb and 1.33 kb band intensities.

Subjects with a pathological C9ORF72 hexanucleotide expansion in both alleles

The described protocol readily identifies patients with repeat expansions within both *C9ORF72* loci. These cases are characterised by the absence of a 1.33 kb band on Southern blots, as illustrated in Figure 2, lane 1 for a patient carrying one allele with ~50 and another allele with >2000 GGGGCC repeats in the *C9ORF72* locus.

Southern hybridization in different tissues from the same subject

When our Southern hybridisation protocol was used for analysis of genomic DNA extracted from peripheral blood, cerebellum, cortex and a lymphoblastoid cell line of patients with GGGGCC repeat expansion, a different pattern of high molecular mass fragments has been detected for each DNA sample of the same patient, as illustrated in Figure 3 and Additional file 2: Figure S1. This observation strongly suggests that an expanded repeat region in the *C9ORF72* locus is unstable not only in nucleated blood cells but also in other somatic cells, including cells of the central nervous system.

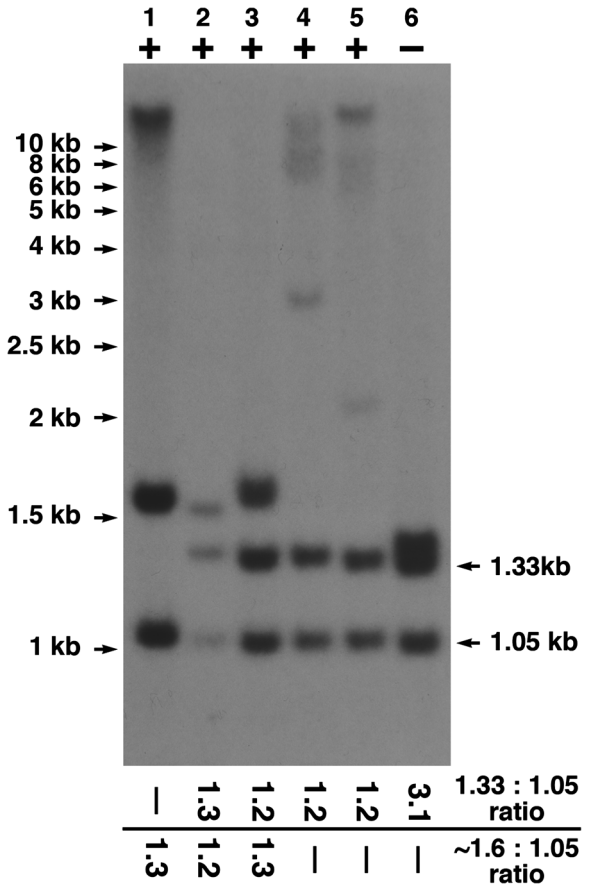


Figure 2 Identification of a patient with GGGGCC repeat expansion in both alleles of the *C9ORF72* locus. Hybridisation of EcoRI and XbaI digested DNA extracted from the cultured lymphoblastoid cell lines of patients positive (+) or negative (–) for GGGGCC repeat expansion according to results of the repeat-PCR analysis. Note the presence of ~1.6 kb band representing an allele with relatively small, although what is considered pathological, i.e. >30, number of repeats in one of the *C9ORF72* loci in the genome of three patients (lanes 1 – 3). In contrast to patients carrying the second allele without pathological repeat expansion (lanes 2–6), one patient (lane 1) has the second allele with very high number of repeats. Where relevant, the ratios of the 1.33 kb to 1.05 kb and/or ~1.6 kb to 1.05 kb bands intensities are shown below the image. Note that for sample 6 the intensity used for calculation is combined intensities of two close ~1.33 kb size bands.

Although a ^{32}P-radiolabeled hybridisation probe remains the best option for detection of GGGGCC repeat expansions in the *C9ORF72* locus with the generation of clearer and more quantitative results compared to those obtained using DIG-labeled probes, in a pilot experiments we

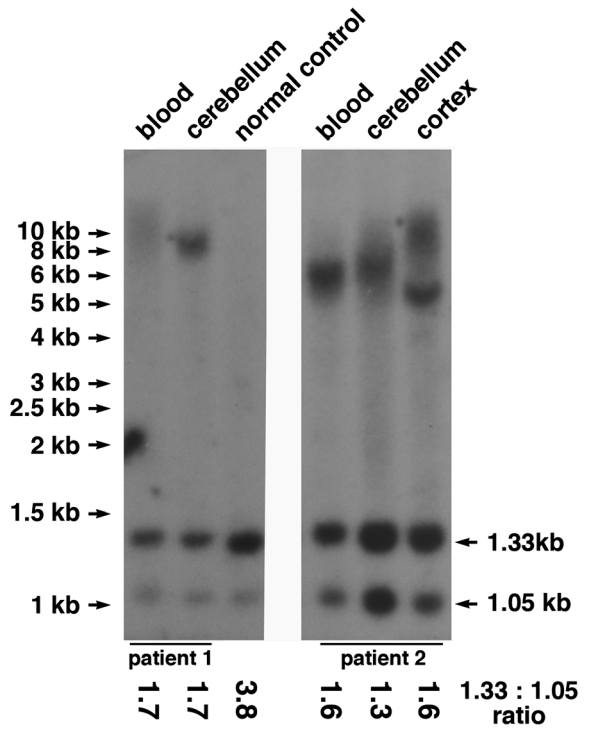


Figure 3 Somatic cell genomes of ALS patients have variable number of GGGGCC repeats in the *C9ORF72* locus. Two sections of the same Southern blot (see raw image in Additional file 2: Figure S1) show hybridisation of EcoRI and XbaI digested DNA extracted from peripheral blood, cerebellum and cerebral cortex of two patients positive for GGGGCC repeat expansion. The ratios of the 1.33 kb to 1.05 kb bands intensities are shown below the image. Note that in this case these ratios are unconventionally high, which is due to the gel has been run for too long and therefore the 1.05 kb bands appeared too close to the edge of the gel, resulting in less efficient transfer and hybridisation of this band. However, the ratio for normal DNA sample is still approximately twice as high as the ratio for all DNA samples with GGGGCC repeat expansion.

successfully used the latter detection method with our protocol (data not shown).

Conclusions

Our optimised protocol for Southern hybridisation allows simple and confident detection, as well as sizing of the repeat expansion in various types of cells and tissues. Thus we recommend this protocol as an acceptable standard for Southern blot hybridisation analysis of GGGGCC repeat expansions in the *C9ORF72* locus.

Methods

DNA extraction

Whole blood samples were obtained from the Sheffield MND Blood DNA Biobank and CNS tissue was obtained from the Sheffield Brain Tissue Bank. The South Sheffield Research Ethics Committee approved the study, and informed consent was obtained for all samples. Lymphoblastoid cell lines were obtained from patients with ALS from the Wellcome Trust/Motor Neurone Disease Association-funded ALS/MND DNA bank and associated lymphoblastoid cell line repository in the UK.

Before lysis with a digest buffer (50 mM Tris–HCl pH 8.0; 100 mM NaCl; 5 mM EDTA; 1% SDS; 2 mg/ml proteinase K) frozen CNS tissues were grinded in liquid nitrogen. Following overnight digestion at 55°C genomic DNA was extracted twice with phenol and once with phenol/chloroform and ethanol precipitated.

Southern hybridisation

A genomic fragment was amplified from a human DNA using oligonucleotide primers AGTTCCAGAGCTTG CTACAG and GAACAGTAGGAAAAGGGTCTG and cloned into the pCR-BluntII-TOPO vector (Invitrogen) to produce a pCh9.1 plasmid carrying an insert used as the hybridisation probe (Additional file 3: Figure S2). Well-established and widely used methods of Southern transfer, preparation of the probe and hybridisation procedure were used. In brief, approximately 20 μg of genomic DNA was digested with EcoRI and XbaI and the resulting fragments separated in 1% TAE agarose gel. After incubating the gel in 0.25 N HCl for 20 min, 0.5 N NaOH; 1.5 M NaCl for 40 min and 0.5 M Tris–HCl pH7.2; 3 M NaCl for 40 min at room temperature, DNA was transferred to a nylon membrane (Hybond N+, GE Healthcare) by capillary blotting [10]. An eukaryotic insert of the plasmid pCh9.1 was excised using KpnI and XbaI, separated from the plasmid backbone in agarose gel, purified using Qiagen kit and used for preparing hybridisation probes. The probe was labelled with ^{32}P in a nick-translation reaction [11]. Blots were prehybridised in HB (4×SSC; 0.5% SDS; 5×Denhardt's solution; 100 μg/ml denatured salmon testis DNA) at 67°C for 4 h. Labelled DNA was denatured by incubation at 100°C for 5 min followed by immediate mixing with ice–cold HB. Hybridisation was carried out at 67°C for 16–20 h. Blots were washed 3 times in in 2×SSC; 0.2% SDS at 67°C and exposed to X-ray film. The detailed protocol as well as the pCh9.1 plasmid carrying an insert used as the hybridisation probe can be obtained by sending a request to Vladimir Buchman (buchmanvl@cf.ac.uk).

DNA extracted from 32 lymphoblastoid cell lines obtained from the MND National Biobank was used for Southern hybridisation analysis. This included 29 cases with a *C9ORF72* expansion identified by repeat-PCR and 3 cases with non-*C9ORF72* ALS. For certain cases DNA was also extracted from peripheral blood and CNS tissue available from the Sheffield Brain Tissue and DNA bank.

Quantification of the band intensity ratios

When at least one normal (without repeat expansion) DNA sample is present on the blot as a reference, the

difference in relative intensity of 1.33 kb and 1.05 kb bands in this sample and in samples with repeat expansion can be easily recognised by eye. However, this difference can be quantified by scanning X-ray films, measuring bands using a software available for every gel documentation system and calculating the ratio of pixels in 1.33 kb and 1.05 bands for each sample. Even when for any technical reasons, the ratio of bands in samples with repeat expansion is higher than conventional 1:1, this is accompanied by corresponding increase of the ratio in a normal sample (s) present on the same Southern blot, which still allows discrimination of repeat-bearing and normal cases by eye. An example of this is shown in Figure 3.

Estimation of the number of repeats within an expanded allele

The number of repeats within the expanded allele is estimated based on the size of corresponding fragment in base pairs minus the size of the non-expanded fragment (1330 base pairs) and divided by the size of the repeat unit (6 base pairs). For example, if the size of a fragment was 3 kb, estimated number of repeats in the corresponding locus is (3000 – 1330)/6 = 278. As with any method based on agarose gel electrophoresis, an accuracy of the repeat number estimate using the described protocol varies with the size of detected fragments. In the range of the repeat-bearing fragment size between 1.5 kb (which corresponds to an upper limit for a "normal" allele, i.e. less than 30 repeats = 0.18 kb + 1.33 kb -> 1.51 kb) and 3 kb (~300 repeats) the accuracy of the estimate might be around 10 repeats; between 3 kb and 6 kb (~750 repeats) it drops to ~50 repeats and between 6 kb and 10 kb (~1500 repeats) – to 200 repeats. Any bands above 12 kb should be considered as >2000 repeats.

Additional files

Additional file 1: Table S1. Ratios of 1.33 kb and 1.05 kb bands intensities for samples shown in Figure 1C.

Additional file 2: Figure S1. Raw image of the Southern blot used for preparing main Figure 3.

Additional file 3: Figure S2. Scheme of a DNA fragment used as a hybridisation probe inserted into a polylinker cloning site of pCR-Blunt II-TOPO plasmid. The sequence of the plasmid polylinker region is from the Invitrogen manual for TOPO cloning kit.

Abbreviations

ALS: Amyotrophic lateral sclerosis; CNS: Central nervous system; DIG: Digoxigenin.

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

VB and NN designed the Southern hybridisation probe and conceived the protocol. JCK, JK and PJS participated in the design of the study. VB, JCK, NN, AH, ODR and NCR carried out experiments. VB, NN, JCK and PJS participated in data analysis. VB, JCK and PJS drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This work was supported by Research Grants from the Welcome Trust (075615/Z/04/z) and RF State Programme (agreement No 8829) to VLB, and EU Framework 7 (Euromotor No259867) to PJS and JK. JCK is supported by an MND Association / Medical Research Council Lady Edith Wolfson Fellowship award (MR/K003771/1). The biosample collection for the ALS cases was supported by the MND Association and the Wellcome Trust (PJS). We are extremely grateful to all of the patients who donated biosamples for research purposes.

Author details

[1]School of Biosciences, Cardiff University, Museum Avenue, Cardiff CF10 3AX, UK. [2]Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, 385A Glossop Road, Sheffield S10 2HQ, UK. [3]Pirogov Russian National Research Medical University, Ostrovitianov str. 1, Moscow 117997, Russian Federation. [4]Institute of Physiologically Active Compounds of RAS, 1 Severniy Proezd, Chernogolovka 142432, Moscow Region, Russian Federation.

Received: 6 February 2013 Accepted: 22 March 2013
Published: 8 April 2013

References

1. DeJesus-Hernandez M, Mackenzie, Ian R, Boeve, Bradley F, Boxer, Adam L, Baker M, Rutherford, Nicola J, Nicholson, Alexandra M, Finch, NiCole A, Flynn H, Adamson J, Kouri N, Wojtas A, Sengdy P, Hsiung G, Yuek R, Karydas A, Seeley, William W, Josephs, Keith A, Coppola G, Geschwind, Daniel H, Wszolek, Zbigniew K, Feldman H, Knopman, David S, Petersen, Ronald C, Miller, Bruce L, Dickson, Dennis W, Boylan, Kevin B, Graff-Radford, Neill R, Rademakers R: **Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS.** *Neuron* 2011, **72:**245–256.
2. Renton, Alan E, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, Schymick, Jennifer C, Laaksovirta H, Van Swieten, John C, Myllykangas L, Kalimo H, Paetau A, Abramzon Y, Remes, Anne M, Kaganovich A, Scholz, Sonja W, Duckworth J, Ding J, Harmer, Daniel W, Hernandez, Dena G, Johnson, Janel O, Mok K, Ryten M, Trabzuni D, Guerreiro, Rita J, Orrell, Richard W, Neal J, Murray A, Pearson J, Jansen, Iris E, *et al*: **A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD.** *Neuron* 2011, **72:**257–268.
3. Ash PE, Bieniek KF, Gendron TF, Caulfield T, Lin WL, DeJesus-Hernandez M, van Blitterswijk MM, Jansen-West K, Paul JW 3rd, Rademakers R, Boylan KB, Dickson DW, Petrucelli L: **Unconventional Translation of C9ORF72 GGGGCC Expansion Generates Insoluble Polypeptides Specific to c9FTD/ALS.** *Neuron* 2013, **77:**639–646.
4. Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E, Schmid B, Kretzschmar HA, Cruts M, Van Broeckhoven C, Haass C, Edbauer D: **The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTLD/ALS.** *Science* 2013. doi:10.1126/science.1232927. Published online 7 February 2013.
5. Todd PK, Paulson HL: **RNA-mediated neurodegeneration in repeat expansion disorders.** *Ann Neurol* 2010, **67:**291–300.
6. Gijselinck I, Van Langenhove T, van der Zee J, Sleegers K, Philtjens S, Kleinberger G, Janssens J, Bettens K, Van Cauwenberghe C, Pereson S, Engelborghs S, Sieben A, De Jonghe P, Vandenberghe R, Santens P, De Bleecker J, Maes G, Bäumer V, Dillen L, Joris G, Cuijt I, Corsmit E, Elinck E, Van Dongen J, Vermeulen S, Van den Broeck M, Vaerenberg C, Mattheijssens M, Peeters K, Robberecht W, *et al*: **A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study.** *Lancet Neurol* 2012, **11:**54–65.
7. Prior TW: **Technical standards and guidelines for myotonic dystrophy type 1 testing.** *Genet Med* 2009, **11:**552–555.
8. Ishiura H, Takahashi Y, Mitsui J, Yoshida S, Kihira T, Kokubo Y, Kuzuhara S, Ranum L, Tamaoki T, Ichikawa Y, Date H, Goto J, Tsuji S: **C9orf72 repeat expansion in amyotrophic lateral sclerosis in the kii peninsula of japan.** *Arch Neurol* 2012, **69:**1154–1158.
9. Takada L, Pimentel M, Dejesus-Hernandez M, Fong J, Yokoyama J, Karydas A, Thibodeau M, Rutherford N, Baker M, Lomen-Hoerth C, Rademakers R, Miller

B: **Frontotemporal dementia in a Brazilian kindred with the c9orf72 mutation.** *Arch Neurol* 2012, **69:**1149–1153.
10. Southern E: **Detection of specific sequences among DNA fragments separated by gel electrophoresis.** *J Mol Biol* 1975, **98:**503–17.
11. Rigby PWJ, Dieckmann M, Rhodes C, Berg P: **Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I.** *J Mol Biol* 1977, **113:**237–251.

doi:10.1186/1750-1326-8-12
Cite this article as: Buchman *et al.*: Simultaneous and independent detection of C9ORF72 alleles with low and high number of GGGGCC repeats using an optimised protocol of Southern blot hybridisation. *Molecular Neurodegeneration* 2013 **8**:12.

Published in final edited form as:
Acta Neuropathol. 2014 March ; 127(3): 407–418. doi:10.1007/s00401-013-1239-x.

TMEM106B is a genetic modifier of frontotemporal lobar degeneration with *C9orf72* hexanucleotide repeat expansions

Michael D. Gallagher[1,2], **Eunran Suh**[3], **Murray Grossman**[2], **Lauren Elman**[2], **Leo McCluskey**[2], **John C. Van Swieten**[4,5], **Safa Al-Sarraj**[6], **Manuela Neumann**[7,8], **Ellen Gelpi**[9], **Bernardino Ghetti**[10], **Jonathan D. Rohrer**[11], **Glenda Halliday**[12,13], **Christine Van Broeckhoven**[14], **Danielle Seilhean**[15], **Pamela J. Shaw**[16], **Matthew P. Frosch**[17], **International Collaboration for Frontotemporal Lobar Degeneration**[†], **John Q. Trojanowski**[3], **Virginia M.Y. Lee**[3], **Vivianna Van Deerlin**[3], and **Alice S. Chen-Plotkin**[2]

[1]Cell & Molecular Biology Graduate Group, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

[2]Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

[3]Center for Neurodegenerative Disease Research, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

[4]Erasmus Medical Centre, s'Gravendijkwal 230, Rotterdam

[5]Alzheimercenter Vumc, Boelelaan 1118, Amsterdam

[6]King's College Hospital, London

[7]University of Tübingen, Calwerstr. 3, 72072 Tübingen, Germany

[8]German Center for Neurodegenerative Diseases (DZNE)

[9]Neurological Tissue Bank of the Biobank-Hospital Clinic-Insitut d'Investigacions Biomèdiques August Pi i Sunyer, Facultad de Medicina, c/Casanova 143, planta 0, ala sur. 08036 Barcelona, Spain

[10]Department of Pathology & Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN

[11]Dementia Research Centre, Department of Neurodegenerative Disease, UCL Institute of Neurology, London, UK

[12]Neuroscience Research Australia, Barker St, Randwick, NSW 2031, Australia

[13]Faculty of Medicine, University of New South Wales, Australia

[14]Neurodegenerative Brain Disease Group, Department of Molecular Genetics, VIB, Universiteitsplein 1, 2610 Antwerpen, Belgium

[15]University Pierre et Marie Curie (UPMC)-Sorbonne University, France

Correspondence to: Alice Chen-Plotkin, Department of Neurology, 3 W Gates, 3400 Spruce St, Philadelphia, PA 19104, chenplot@mail.med.upenn.edu, Telephone: 215-573-7193, Fax: 215-349-5579.
[†]see International Collaboration for Frontotemporal Lobar Degeneration section for full list of contributors

[16]University of Sheffield, UK

[17]Massachusetts Alzheimer's Disease Research Center, Harvard Medical School, Boston, MA

Abstract

Hexanucleotide repeat expansions in chromosome 9 open reading frame 72 (*C9orf72)* have recently been linked to frontotemporal lobar degeneration (FTLD) and amyotrophic lateral sclerosis (ALS), and may be the most common genetic cause of both neurodegenerative diseases. Genetic variants at *TMEM106B* influence risk for the most common neuropathological subtype of FTLD, characterized by inclusions of TAR DNA binding protein of 43kDa (FTLD-TDP). Previous reports have shown that *TMEM106B* is a genetic modifier of FTLD-TDP caused by progranulin (*GRN*) mutations, with the major (risk) allele of rs1990622 associating with earlier age at onset of disease. Here we report that rs1990622 genotype affects age at death in a single-site discovery cohort of FTLD patients with *C9orf72* expansions (n=14), with the major allele correlated with later age at death (p=0.024). We replicate this modifier effect in a 30-site international neuropathological cohort of FTLD-TDP patients with *C9orf72* expansions (n=75), again finding that the major allele associates with later age at death (p=0.016), as well as later age at onset (p=0.019). In contrast, *TMEM106B* genotype does not affect age at onset or death in 241 FTLD-TDP cases negative for *GRN* mutations or *C9orf72* expansions. Thus, *TMEM106B* is a genetic modifier of FTLD with *C9orf72* expansions. Intriguingly, the genotype that confers increased risk for developing FTLD-TDP (major, or T, allele of rs1990622) is associated with later age at onset and death in *C9orf72* expansion carriers, providing an example of sign epistasis in human neurodegenerative disease.

Keywords

TMEM106B; C9orf72; frontotemporal dementia; frontotemporal lobar degeneration; amyotrophic lateral sclerosis; genetic modifier

INTRODUCTION

Frontotemporal lobar degeneration (FTLD) is the second most common dementia in individuals under 65 years of age [30]. The most common neuropathological subtype is frontotemporal lobar degeneration with TAR DNA-binding protein of 43kDa (TDP-43) inclusions (FTLD-TDP) [30]. We previously reported the minimally characterized gene, *TMEM106B*, as a risk factor for FTLD-TDP by genome-wide association study (GWAS) [38], and this association has been verified independently [12,39]. In our GWAS, three SNPs reached genome-wide significance for association with FTLD-TDP [38]; all are located within a 36kb haplotype block that contains *TMEM106B* and no other genes. The major alleles of all three SNPs are associated with increased risk of FTLD-TDP ($p=1.08\times10^{-11}$, odds ratio=1.64 for major allele of rs1990622, the top GWAS SNP) [38].

Several studies have begun to elucidate the role *TMEM106B* plays in FTLD-TDP. TMEM106B levels have been shown to be increased in FTLD-TDP brains [5,38], and risk-associated alleles resulting in amino acid variation in the TMEM106B protein have been reported to result in higher steady-state levels of TMEM106B through slower protein

degradation [26]. In addition, the major allele of rs1990622 has been associated with reduced plasma progranulin (PGRN) levels in both healthy individuals and in individuals with FTLD-TDP caused by mutations in *GRN*, the gene encoding progranulin [9,12]. Mutations in *GRN* are a major cause of familial FTLD-TDP [14], and are thought to cause disease via haploinsufficiency of the progranulin protein [14,31]. Interestingly, among *GRN* mutation carriers with FTLD (*GRN(+)* FTLD), *TMEM106B* rs1990622 major alleles have been reported to associate with earlier age at disease onset [9]. Experiments in cell culture systems have also demonstrated that TMEM106B and PGRN co-localize in several cell types, including neurons, and that over-expression of *TMEM106B* alters intra- and extracellular levels of PGRN [3,5,26]. Therefore, increased expression of *TMEM106B* may confer risk for FTLD-TDP by altering PGRN levels.

While *GRN* mutations account for ~5% of clinical FTLD cases [14], and other rarer, monogenic causes of FTLD are known (including mutations in *MAPT*, *CHMP2B* and *VCP*) [17,33,41], a substantial proportion of familial cases were until recently of unknown cause. This changed in late 2011 when two groups reported that hexanucleotide repeat expansions in the *C9orf72* gene are perhaps the most common cause of familial FTLD, familial amyotrophic lateral sclerosis (ALS), and familial FTLD with motor neuron disease (FTLD-MND) [11,28]. Although these mutations display an autosomal dominant mode of inheritance, 3–6% of apparently sporadic cases of FTLD and ALS harbor *C9orf72* expansions as well, which may be explained by genetic anticipation, *de novo* mutation, or incomplete penetrance [11,28].

The function(s) of *C9orf72* and its role in disease are currently areas of ongoing research [10], with evidence for both loss-of-function [8,11,15,28] and gain-of-toxic-function [1,13,25] mechanisms. At a neuropathological level, *C9orf72* expansion positive FTLD (*C9orf72(+)* FTLD) and ALS (*C9orf72(+)* ALS) cases exhibit TDP-43 pathology reminiscent of *GRN(+)* FTLD, as well as mutation-negative ALS and FTLD, although *C9orf72(+)* FTLD and ALS cases show unique pathological features as well [2,34,35].

Here, we assess whether *TMEM106B* risk genotypes exert a genetic modifier effect in *C9orf72(+)* FTLD and ALS, *GRN(+)* FTLD, and FTLD cases without either mutation. We also investigate whether these genotypes are associated with disease status in *C9orf72(+)* FTLD and with plasma progranulin levels in *C9orf72(+)* expansion carriers.

METHODS

Patient cohorts

FTLD and ALS cases with *C9orf72* expansions of greater than 30 hexanucleotide repeats were identified from among cases in the Integrated Neurodegenerative Disease Database at the University of Pennsylvania (UPenn) to form a discovery cohort [37,44]. Patients were initially seen at the UPenn Frontotemporal Degeneration Center (FTDC), Amyotrophic Lateral Sclerosis Center (ALSC), or Alzheimer's Disease Center (ADC); all were collected with Institutional Review Board Approval. In addition to having a *C9orf72* expansion, the criteria for selection of FTLD cases was a pathological diagnosis of FTLD-TDP (n=10) or a clinical diagnosis of FTLD or FTLD-MND (n=19), according to published criteria [16,22–

24,27,36]. *C9orf72(+)* ALS cases (n=55) all met El Escorial-revised criteria [4]. Twenty of the 55 ALS cases had autopsy confirmation of ALS pathology. For both FTLD and ALS cases, only probands were selected. In situations where patients exhibited both dementia and motor neuron disease (MND), cases were assigned to FTLD-MND if the initial presentation was cognitive and to ALS if the initial presentation was MND. All *C9orf72(+)* FTLD and *C9orf72(+)* ALS cases meeting these criteria were included without bias for familial-vs.-apparently-sporadic patterns of inheritance, and without prior knowledge of *TMEM106B* genotype.

The *C9orf72(+)* FTLD discovery cohort is 93.5% white (6.5% unknown ethnicity) and 54.8% male. The *C9orf72(+)* ALS cohort is 87.2% white, 5.6% black, 3.5% Latino, and 3.7% unknown ethnicity with 59.8% males. Age at onset and age at death were collected, but both were not available on all subjects (*e.g.* no age at death for living subjects, and sometimes no known age at onset for autopsy cases), therefore the numbers of cases from each cohort vary depending on the data needed for analysis. For the discovery cohort, age at onset was defined as the age at initial complaint, based on review of medical records.

The previously published and publicly available FTLD-TDP GWAS from the International Collaboration for Frontotemporal Lobar Degeneration was used as a replication cohort [38]. As previously described [38], all cases of this postmortem cohort were self-described as White, of European ancestry. In addition, samples were screened by principle components analysis of genomewide genotyping data, and at >200 ancestry informative markers, to reduce effects of population stratification. Only those cases with >90% inferred CEU (based on HapMap CEU population of Utah residents with ancestry from Northern and Western Europe) ancestry were included in the original GWAS [38], from which all cases of the current replication cohort are derived.

A subset of the FTLD-TDP cases were known from the original study to have a pathogenic *GRN* mutation (n=116) and are used here as a comparison group [7,38]. The majority of cases lacking a *GRN* or *VCP* mutation (n=321) were screened for *C9orf72* expansions either by the contributing site or by UPenn, using published methods [11,28]. 80 FTLD-TDP cases with *C9orf72* expansions were identified from 30 clinical sites that agreed to collaborate on this project (see Acknowledgement section for a full listing of clinical sites). Of the 80 cases, 5 UPenn cases overlapped with the UPenn discovery cohort and were removed, leaving 75 *C9orf72* expansion cases for analysis in the replication cohort. In addition, 241 cases were formally tested for (and found negative for) *C9orf72* expansions, and these were used as the mutation-negative FTLD-TDP cohort. We note that there were additional *C9orf72(+)* FTLD-TDP cases in the GWAS, but only those cases from sites agreeing to collaborate on this study (constituting >80% of the total FTLD-TDP GWAS *C9orf72(+)* cases) are included here.

For the replication cohort, age at onset and age at death were provided by the contributing clinical site.

Genotyping

DNA from UPenn cases, extracted from blood or brain samples as previously described [38], was tested for rs1990622 genotype using one of two methods: TaqMan chemistry-based allelic discrimination assays as previously described [5,38], or a custom Sequenom MassArray genotyping panel that includes PCR and extension primers for rs1990622. PCR and extension primer sequences for the Sequenom panel are available on request. Both genotyping methods were compared and found to be concordant (data not shown) [37].

Plasma progranulin measurement

Plasma samples were collected from UPenn ALS and FTLD discovery cohort patients, aliquotted, and stored at −80°C as previously described [6]. Progranulin levels were measured using a commercially available sandwich ELISA (Human progranulin ELISA kit, AdipoGen), according to manufacturer instructions.

Statistical analyses

Linear regression analyses evaluating the association of *TMEM106B* genotype with age at death or age at disease onset were performed in R, with or without covariates as described in the text. Two-tailed p-values are reported for the discovery cohort, and one-tailed p-values are reported for the FTLD-TDP GWAS replication cohort, since the expected directionality was known. For the combined dataset, survival analyses (Kaplan-Meier method) were also performed in Prism, and two-tailed p-values from the log-rank test for trend are reported.

Where indicated, codominant, major-allele-dominant, and minor-allele dominant models of genetic effect were investigated.

In addition, we tested for association between *TMEM106B* genotype and disease for genetically-defined subsets of FTLD (*C9orf72(+)* FTLD, *GRN(+)* FTLD, or individuals without *C9orf72* expansions or *GRN* mutations). Chi-square statistics were calculated for rs1990622 using the FTLD-TDP GWAS cases and controls [38].

For plasma progranulin analyses, Kruskal-Wallis tests were used to compare plasma progranulin measures among carriers of different *TMEM106B* genotypes under a codominant model, and Mann-Whitney tests were used to compare different *TMEM106B* genotypes under major-allele-dominant and minor-allele dominant models. In addition, multivariate linear regressions predicting plasma progranulin levels from *TMEM106B* genotype were used to adjust for sex, age, duration of disease, or clinical manifestation as described in the text.

R-scripts for analyses are available upon request.

RESULTS

TMEM106B genotype at rs1990622 influences age at death in a discovery cohort of *C9orf72(+)* FTLD

TMEM106B genotype has been shown to demonstrate a genetic modifier effect in FTLD-TDP caused by autosomal dominant mutations in the progranulin gene (*GRN*) [9]. We

therefore asked whether genetic variation at *TMEM106B* influences age at death or age at onset in *C9orf72(+)* FTLD or ALS disease cases. We assumed a codominant model for these initial analyses.

In *C9orf72(+)* FTLD (n=14), age at death was significantly correlated with *TMEM106B* genotype at rs1990622, the SNP previously found in our GWAS to associate most strongly with FTLD-TDP risk (p=0.024, Table 1). Adjusting for sex and presence/absence of co-existing MND did not affect this association. Moreover, the direction of association was surprising; specifically, the major allele of rs1990622 (C) was associated with later age at death in *C9orf72(+)* FTLD. In our GWAS, the major allele of rs1990622 was found to be associated with increased risk for the development of FTLD.

In contrast, rs1990622 genotype did not affect age at death in *C9orf72(+)* ALS (n=39, Table 1). In this discovery cohort, rs1990622 genotype did not affect age at onset for *C9orf72* expansion carriers who presented with either ALS (n=47) or FTLD (n=26). However, a statistically significant association emerged when we performed a multivariate analysis controlling for gender and presence of FTD in the clinical ALS cases, with the major allele associating with earlier age at onset (n=47, Table 1).

TMEM106B genotype at rs1990622 influences age at onset and age at death in a replication cohort of *C9orf72(+)* FTLD

We sought to replicate the genetic modifier effect of *TMEM106B* in *C9orf72(+)* FTLD in an independent cohort of patients. Since the majority of cases from our GWAS had been screened for the presence of *C9orf72* expansions, these cases provided an ideal replication cohort to evaluate the effect of *TMEM106B* rs1990622 genotype on age at death in *C9orf72(+)* FTLD for three key reasons. First, since the FTLD-TDP GWAS predated the discovery of *C9orf72* expansions as a cause of FTLD, this large, international cohort was unbiased in enrollment with respect to *C9orf72* status. Second, all cases were neuropathologically confirmed to have FTLD-TDP, ensuring neuropathological homogeneity. Third, because all cases had undergone genome-wide genotyping and filtering for effects from population stratification, we could be certain that effects from cryptic familial relationships or population stratification would be minimal.

As shown in Table 2, rs1990622 genotype was again correlated with age at death in this cohort (n=75), in both univariate analyses (p=0.016) and linear regression models adjusting for sex and the presence or absence of MND (p=0.019). Moreover, in this larger replication cohort, rs1990622 genotype was also correlated with age at onset (n=68 with age at onset data, p=0.019 for univariate analyses and p=0.032 for multivariate analyses adjusting for sex and presence or absence of MND). Consistent with the results from our discovery cohort, the major allele (T) of rs1990622 was associated with later age at death, as well as later age at onset. Indeed, patients showed later disease onset and later death by more than three years for each additional major allele at rs1990622 carried.

We further examined this genetic modifier effect using Kaplan-Meier survival analyses performed on the combined cohort (discovery plus replication, n=89 for age at death analysis, n=94 for age at onset analysis) of *C9orf72(+)* FTLD cases. As shown in Fig. 1,

TMEM106B genotypes at rs1990622 were significantly associated with age at death (Fig. 1A, p=0.046, log rank test for trend), with a trend towards association for age at onset (Fig. 1C, p=0.064) in this combined cohort. In addition, we observed that the curve separation between rs1990622 minor allele homozygotes (CC) and heterozygotes (TC) was greater than the separation between heterozygotes (TC) and major allele homozygotes (TT). We therefore re-analyzed our data under a major-allele dominant model for rs1990622 and observed a stronger effect of *TMEM106B* genotype on age at death (p=0.041, log rank test for trend) and age at onset (p=0.037, log rank test for trend) in *C9orf72(+)* FTLD. Indeed, at any given age, minor allele (C) homozygotes at rs1990622 had more than twice the risk of manifesting disease (Fig. 1D, HR 2.022, 95% CI 1.042–3.925), and more than twice the risk of death (Fig. 1B, HR 2.039, 95% CI 1.031–4.033), compared to other genotypes.

TMEM106B genotype does not exert a genetic modifier effect in *C9orf72* expansion negative FTLD-TDP cases

We next asked whether the *TMEM106B* genetic modifier effect observed for *C9orf72(+)* FTLD extended to FTLD-TDP cases without *C9orf72* expansions, again using FTLD-TDP cases from the FTLD-TDP GWAS for which *C9orf72* and/or *GRN* mutation status was known. We considered cases with and without *GRN* mutations separately.

As shown in Fig. 2A, *TMEM106B* rs1990622 genotype did not affect age at death in FTLD-TDP cases without *C9orf72* expansions or *GRN* mutations (n=241). In the subset of *GRN*-related FTLD-TDP (n=116, Fig. 2B), only one rs1990622 CC individual had age at death information available, so we could only compare TT and TC individuals, who did not differ significantly in age at death. Similar results were obtained for age-at-onset analyses (data not shown).

TMEM106B genotype is associated with FTLD-TDP in *C9orf72* expansion carriers

The observed genetic modifier effect for *TMEM106B* in *C9orf72(+)* FTLD is surprising in its direction. Specifically, the rs1990622 major allele associated with increased risk of FTLD-TDP by GWAS is correlated with older age at onset and death among *C9orf72(+)* FTLD cases, implying a beneficial effect in this mutation subgroup. We therefore examined *TMEM106B* rs1990622 allele frequencies in 116 *GRN(+)* FTLD cases, 80 *C9orf72(+)* FTLD cases, and 241 FTLD-TDP cases in which mutations in *GRN* and expansions in *C9orf72* had been excluded. As with the age-at-onset and age-at-death analyses, FTLD-TDP cases were from our prior FTLD-TDP GWAS, although numbers in each group are slightly higher because individuals with genotypes but lacking age-at-death or age-at-onset data could be included. As shown in Table 3, *TMEM106B* rs1990622 genotype was significantly associated with FTLD-TDP in all three subgroups, with the same direction of association in all three subgroups. In each case, the major allele of rs1990622 was enriched in disease.

TMEM106B genotype is not associated with plasma progranulin levels in *C9orf72* expansion carriers

TMEM106B genotype has been reported to influence plasma progranulin levels in healthy individuals and *GRN*+ FTLD, with the rs1990622 major allele associated with decreased progranulin expression. We evaluated whether this relationship was also true in *C9orf72*

expansion carriers. In a convenience subset of 24 *C9orf72* expansion carriers (20 with *C9orf72(+)* ALS and 4 with *C9orf72(+)* FTLD) from the UPenn discovery cohort for whom we had plasma samples, we measured progranulin levels using an enzyme-linked immunosorbent assay (ELISA). As shown in Fig. 2C, there were no significant differences in plasma progranulin levels comparing *C9orf72* expansion carriers with TT, TC, and CC genotypes at rs1990622. Adjusting for sex and age at plasma sampling or duration of disease did not affect this result. Additionally adjusting for clinical manifestation as FTLD or ALS did not affect this result.

DISCUSSION

In the current study, we find that *TMEM106B* is a genetic modifier for *C9orf72(+)* FTLD, demonstrating a significantly later age at death and age at onset for *TMEM106B* rs1990622 major allele (T) carriers. This effect appears to be specific to *C9orf72(+)* FTLD, since *C9orf72*(−)FTLD cases do not differ in age at death depending on rs1990622 genotype. In addition, rs1990622 major allele carriers are significantly enriched in *C9orf72(+)* FTLD, compared to neurologically normal controls. Finally, among *C9orf72* expansion carriers, we do not see a clear effect of rs1990622 genotype on plasma progranulin levels.

We observe that *TMEM106B* genotypes exert a genetic modifier effect in *C9orf72(+)* FTLD. Examples of common risk variants acting as genetic modifiers in Mendelian subgroups of disease are increasingly being described. In the field of neurodegeneration, one well-known example is the age-at-onset modifying effect of Apolipoprotein E (*APOE*) isoform in *PSEN2*-related-Alzheimer's Disease [43]. Moreover, in *GRN*+ FTLD, *TMEM106B* has been reported as a genetic modifier affecting both age-at-onset and circulating levels of progranulin [9,12].

What is more unusual in this case is the direction of the genetic modifier effect. Specifically, the *TMEM106B* allele that is associated with increased risk of developing FTLD-TDP [38] (and earlier age at onset in *GRN*+ FTLD [9]) appears to ameliorate the disease phenotype (associating with later age at death and onset) in *C9orf72(+)* FTLD. This effect may be an example of the general phenomenon of sign epistasis, in which a genetic variant is beneficial on some genetic backgrounds but deleterious in others. In this case, the genetic variant in question is *TMEM106B* genotype at rs1990622 (and linked SNPs), and the genetic backgrounds demonstrating opposing effects are (1) *C9orf72*(+) individuals -- where the major allele at rs1990622 and linked SNPs is protective in modulating the severity of FTLD manifestation, as demonstrated by older age at onset and age at death and (2) *C9orf72*(−) individuals -- where the major allele at rs1990622 and linked SNPs is harmful in conferring increased risk of developing FTLD.

Sign epistasis has its conceptual underpinnings in the evolutionary biology literature [42]. With the advent of modern experimental tools, sign epistasis has been demonstrated in lower organisms such as bacteria [32], with reports for this phenomenon in the realm of human genetics and human disease genetics as well [18,19]. In the few reported empirically-derived examples of sign epistasis, the two (or more) genetic loci involved converge mechanistically in, for example, antibiotic resistance pathways [29] or enzyme-substrate interactions [45].

Thus, the observed epistasis between *TMEM106B* and *C9orf72* suggests that these two proteins may have convergent functions in the pathophysiology of FTLD-TDP. Intriguingly, TMEM106B has been linked to endosomal-lysosomal pathways [3,5,20,26]. The largely uncharacterized protein C9orf72 is structurally related to DENN protein family members [21]. DENN proteins function in the regulation of Rab GTPases, which in turn regulate the many membrane trafficking events needed for proper function of the endosomal-lysosomal pathway.

We note that *TMEM106B* rs1990622 genotypes differ in allelic frequencies between *C9orf72(+)* FTLD-TDP and normal controls; this situation in which a common variant shows allelic association with disease even in a monogenic, highly-penetrant subgroup of disease has been reported in *GRN*+ FTLD-TDP as well [12,38]. In the case of the *GRN* mutants, a potential explanation may lie in ascertainment bias, since *TMEM106B* risk variant carriers may manifest disease at an earlier age [9], making it more likely for them to be included in a cross-sectional sampling of diseased individuals. Alternately, the protective effect of the modifier locus (*e.g. TMEM106B*) may be significant enough to counter-act the disease-causing effects of the Mendelian genetic cause (*e.g. GRN*), such that carriers of protective variants never manifest clinically despite possessing a highly-penetrant genetic mutation. Such an argument cannot explain our current result, however, since the rs1990622 major allele (found by genome-wide association to be enriched in FTLD-TDP) appears to delay age at death and age at onset in *C9orf72(+)* FTLD cases. An alternate explanation may lie in the fact that *C9orf72* expansions have a broad range of phenotypic expression, manifesting as ALS, FTLD, or a syndrome combining both motor neuron disease and dementia. We have previously shown that ALS patients who are major allele carriers at rs1990622 are more likely to demonstrate cognitive impairment [40]. Thus, it is possible that *TMEM106B* genotype modulates the phenotypic expression of *C9orf72* expansions, with rs1990622 major allele carriers more likely to manifest clinically with dementia. Whether an effect of directing regional pathology towards cognitive regions rather than motor regions also underlies the apparently protective effect on age at death for *TMEM106B* rs1990622 major allele carriers with *C9orf72* expansions remains to be seen.

It is notable that we were able to replicate the genetic modifier effect of *TMEM106B* genotype in *C9orf72(+)* FTLD in a 30-site, international cohort of subjects. Undoubtedly, site-to-site variation in methods of ascertaining age at onset would contribute to noise, and site-to-site variation in practice with respect to aggressiveness of clinical care with a fatal neurodegenerative disease would contribute to differences in age at death in such a dataset. The ability to see a significant genetic modifier effect of *TMEM106B* on *C9orf72* in such a cohort, nonetheless, may have been helped by the fact that our replication cohort was homogeneous with respect to neuropathology (all FTLD-TDP), and genome-wide genotyping in these individuals allowed us to exclude important potential sources of noise, such as population stratification and cryptic familial relationships among individuals. In any case, the international, multi-site nature of our replication cohort increases our confidence that our findings are not due to artifact.

The current study has several limitations. First, while we did not see an age-at-death-modifying effect for *TMEM106B* in *C9orf72* expansion-associated ALS, our sample size

was small (n=39) and likely underpowered to adequately address this question. Thus, future studies examining this relationship in more *C9orf72*-expansion-related ALS cases would be a valuable addition to the data presented here. Second, we did not see a clear modifier effect of *TMEM106B* genotype in the *GRN*(+) FTLD-TDP cases in this study, as has been previously reported [9]. However, our study had only one rs1990622 minor allele homozygote in the *GRN*+ FTLD subgroup, precluding our ability to examine *TMEM106B* genotype effect in a major-allele-dominant model. Third, we were able to obtain plasma samples on 24 *C9orf72* expansion carriers, in whom we measured progranulin levels. Plasma progranulin levels did not differ by *TMEM106B* genotype in this set of samples, which could reflect either insufficient sample size or a biologically-relevant finding. Should further studies in larger sample sizes corroborate our result, this would suggest that *C9orf72* expansions may interrupt the means by which TMEM106B affects circulating progranulin levels. Finally, our study was a targeted evaluation of one locus (*TMEM106B*) for genetic modifier effect in *C9orf72* expansion carriers, rather than a comprehensive screen for genetic modifiers in *C9orf72(+)* FTLD or ALS. It is entirely possible that other loci with epistatic effects exist and also play an important role in modulating the phenotype associated with *C9orf72* expansions. In conclusion, we demonstrate here that *TMEM106B* is the first reported genetic modifier in *C9orf72* expansion-related FTLD. Our findings suggest a previously unsuspected link between these two proteins in the pathophysiology of FTLD and open up new directions for the development of disease-modifying therapy

Acknowledgments

FUNDING

Contributing sites that provided *C9orf72* genetic data included: Erasmus University, Rotterdam, The Netherlands; Indiana University, Indianapolis, Indiana; Banc de Teixits Neurologics-Biobanc-Hospital Clinic-IDIBAPS, Barcelona, Spain; Kings College, London, UK; UCL Institute of Neurology, Queen Square, London, UK; Ludwig-Maximilians University, Munich, Germany; University of New South Wales, Sydney, Australia; VIB, University of Antwerp, Antwerp, Belgium; Massachusetts General Hospital, Boston, Massachusetts; University of Sheffield, Sheffield, UK; Institut National de la Santé et de la Recherche Laboratoire de Neuropathologie, Paris, France.

Contributing sites with *C9orf72(+)* cases identified at UPenn included: Sydney Brain Bank, Australia; Boston University, Boston, Massachusetts; Duke University, Durham, North Carolina; Emory University, Atlanta; Georgia; Karolinska Institute, Stockholm, Sweden; Mt. Sinai School of Medicine, Bronx, New York; Oregon Health Sciences University, Portland, Oregon; University of Pittsburgh, Pittsburgh, Pennsylvania; Rush University, Chicago, Illinois; University of Texas Southwestern, Dallas, Texas; University of Toronto, Toronto, Canada; University of California (Davis, Irvine, San Diego campuses), California; University of Michigan, Ann Arbor, Michigan; University of Kuopio, Finland; University of Southern California, Los Angeles, California; Washington University, St. Louis, Missouri; University of Pennsylvania, Philadelphia, Pennsylvania.

Sources of support for this project include the NIH (AG033101, NS082265, P50 AG005133), The Neurological Tissue Bank of the Biobanc-HC-IDIBAPS, Hersenstichting project BG2010.02, Alzheimer Nederland/NIBC 056-13-018, Stichting Dioraphte projectnr 0802100, The National Institute for Health Research, SOPHIA, EuroMotor, National Health and Medical Research Council of Australia (NHMRC) (FTLD cases supported by NHMRC program grant 1037746), and Neuroscience Research Australia, University of New South Wales. The Antwerp site is in part funded by the MetLife Foundation, USA; the Interuniversity Attraction Poles program of the Belgian Science Policy Office (BELSPO), the Europe Initiative on Centers of Excellence in Neurodegeneration (CoEN) and the Methusalem program supported by the Flemish Government; the Foundation Alzheimer Research (SAO/FRA); the Medical Foundation Queen Elisabeth; the Research Foundation Flanders (FWO); the Agency for Innovation by Science and Technology Flanders (IWT), the University of Antwerp Research Fund, Belgium. The FWO provided a postdoctoral fellowship to J.v.d.Z and I.G. Alice Chen-Plotkin is also supported by the Burroughs Wellcome Fund Career Award for Medical Scientists, a Doris Duke Clinician Scientist Development Award, and the Benaroya Fund. Glenda Halliday holds a NHMRC Senior Principal Research Fellowship. Jonathan D. Rohrer and Martin Rosser are supported by the NIHR Queen Square Dementia Biomedical Research unit and work at the

UCL Institute of Neurology Dementia Research Centre which is supported by Alzheimer's Research UK, Brain Research Trust, and The Wolfson Foundation.

We thank Travis Unger and Beth McCarty Wood for technical assistance. We thank our patients and their families for their participation in this research.

References

1. Ash PE, Bieniek KF, Gendron TF, Caulfield T, Lin WL, Dejesus-Hernandez M, van Blitterswijk MM, Jansen-West K, Paul JW 3rd, Rademakers R, Boylan KB, Dickson DW, Petrucelli L. Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron. 2013; 77:639–646.10.1016/j.neuron.2013.02.004 [PubMed: 23415312]
2. Boxer AL, Mackenzie IR, Boeve BF, Baker M, Seeley WW, Crook R, Feldman H, Hsiung GY, Rutherford N, Laluz V, Whitwell J, Foti D, McDade E, Molano J, Karydas A, Wojtas A, Goldman J, Mirsky J, Sengdy P, Dearmond S, Miller BL, Rademakers R. Clinical, neuroimaging and neuropathological features of a new chromosome 9p-linked FTD-ALS family. J Neurol Neurosurg Psychiatry. 2011; 82:196–203.10.1136/jnnp.2009.204081 [PubMed: 20562461]
3. Brady OA, Zheng Y, Murphy K, Huang M, Hu F. The frontotemporal lobar degeneration risk factor, TMEM106B, regulates lysosomal morphology and function. Hum Mol Genet. 2013; 22:685–695.10.1093/hmg/dds475 [PubMed: 23136129]
4. Brooks BR, Miller RG, Swash M, Munsat TL. World Federation of Neurology Research Group on Motor Neuron Diseases . El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotroph Lateral Scler Other Motor Neuron Disord. 2000; 1:293–299. [PubMed: 11464847]
5. Chen-Plotkin AS, Unger TL, Gallagher MD, Bill E, Kwong LK, Volpicelli-Daley L, Busch JI, Akle S, Grossman M, Van Deerlin V, Trojanowski JQ, Lee VM. TMEM106B, the risk gene for frontotemporal dementia, is regulated by the microRNA-132/212 cluster and affects progranulin pathways. J Neurosci. 2012; 32:11213–11227.10.1523/JNEUROSCI.0521-12.2012 [PubMed: 22895706]
6. Chen-Plotkin AS, Hu WT, Siderowf A, Weintraub D, Goldmann Gross R, Hurtig HI, Xie SX, Arnold SE, Grossman M, Clark CM, Shaw LM, McCluskey L, Elman L, Van Deerlin VM, Lee VM, Soares H, Trojanowski JQ. Plasma epidermal growth factor levels predict cognitive decline in Parkinson disease. Ann Neurol. 2011; 69:655–663.10.1002/ana.22271 [PubMed: 21520231]
7. Chen-Plotkin AS, Martinez-Lage M, Sleiman PM, Hu W, Greene R, Wood EM, Bing S, Grossman M, Schellenberg GD, Hatanpaa KJ, Weiner MF, White CL 3rd, Brooks WS, Halliday GM, Kril JJ, Gearing M, Beach TG, Graff-Radford NR, Dickson DW, Rademakers R, Boeve BF, Pickering-Brown SM, Snowden J, van Swieten JC, Heutink P, Seelaar H, Murrell JR, Ghetti B, Spina S, Grafman J, Kaye JA, Woltjer RL, Mesulam M, Bigio E, Llado A, Miller BL, Alzualde A, Moreno F, Rohrer JD, Mackenzie IR, Feldman HH, Hamilton RL, Cruts M, Engelborghs S, De Deyn PP, Van Broeckhoven C, Bird TD, Cairns NJ, Goate A, Frosch MP, Riederer PF, Bogdanovic N, Lee VM, Trojanowski JQ, Van Deerlin VM. Genetic and clinical features of progranulin-associated frontotemporal lobar degeneration. Arch Neurol. 2011; 68:488–497.10.1001/archneurol.2011.53 [PubMed: 21482928]
8. Ciura S, Lattante S, Le Ber I, Latouche M, Tostivint H, Brice A, Kabashi E. Loss of function of C9orf72 causes motor deficits in a zebrafish model of Amyotrophic Lateral Sclerosis. Ann Neurol. 201310.1002/ana.23946
9. Cruchaga C, Graff C, Chiang HH, Wang J, Hinrichs AL, Spiegel N, Bertelsen S, Mayo K, Norton JB, Morris JC, Goate A. Association of TMEM106B gene polymorphism with age at onset in granulin mutation carriers and plasma granulin protein levels. Arch Neurol. 2011; 68:581–586.10.1001/archneurol.2010.350 [PubMed: 21220649]
10. Cruts M, Gijselinck I, Van Langenhove T, van der Zee J, Van Broeckhoven C. Current insights into the C9orf72 repeat expansion diseases of the FTLD/ALS spectrum. Trends Neurosci. 201310.1016/j.tins.2013.04.010
11. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J, Kouri N, Wojtas A, Sengdy P, Hsiung GY, Karydas A,

Seeley WW, Josephs KA, Coppola G, Geschwind DH, Wszolek ZK, Feldman H, Knopman DS, Petersen RC, Miller BL, Dickson DW, Boylan KB, Graff-Radford NR, Rademakers R. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron. 2011; 72:245–256.10.1016/j.neuron.2011.09.011 [PubMed: 21944778]

12. Finch N, Carrasquillo MM, Baker M, Rutherford NJ, Coppola G, Dejesus-Hernandez M, Crook R, Hunter T, Ghidoni R, Benussi L, Crook J, Finger E, Hantanpaa KJ, Karydas AM, Sengdy P, Gonzalez J, Seeley WW, Johnson N, Beach TG, Mesulam M, Forloni G, Kertesz A, Knopman DS, Uitti R, White CL 3rd, Caselli R, Lippa C, Bigio EH, Wszolek ZK, Binetti G, Mackenzie IR, Miller BL, Boeve BF, Younkin SG, Dickson DW, Petersen RC, Graff-Radford NR, Geschwind DH, Rademakers R. TMEM106B regulates progranulin levels and the penetrance of FTLD in GRN mutation carriers. Neurology. 2011; 76:467–474.10.1212/WNL.0b013e31820a0e3b [PubMed: 21178100]
13. Fratta P, Mizielinska S, Nicoll AJ, Zloh M, Fisher EM, Parkinson G, Isaacs AM. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. Sci Rep. 2012; 2:1016.10.1038/srep01016 [PubMed: 23264878]
14. Gass J, Cannon A, Mackenzie IR, Boeve B, Baker M, Adamson J, Crook R, Melquist S, Kuntz K, Petersen R, Josephs K, Pickering-Brown SM, Graff-Radford N, Uitti R, Dickson D, Wszolek Z, Gonzalez J, Beach TG, Bigio E, Johnson N, Weintraub S, Mesulam M, White CL 3rd, Woodruff B, Caselli R, Hsiung GY, Feldman H, Knopman D, Hutton M, Rademakers R. Mutations in progranulin are a major cause of ubiquitin-positive frontotemporal lobar degeneration. Hum Mol Genet. 2006; 15:2988–3001.10.1093/hmg/ddl241 [PubMed: 16950801]
15. Gijselinck I, Van Langenhove T, van der Zee J, Sleegers K, Philtjens S, Kleinberger G, Janssens J, Bettens K, Van Cauwenberghe C, Pereson S, Engelborghs S, Sieben A, De Jonghe P, Vandenberghe R, Santens P, De Bleecker J, Maes G, Baumer V, Dillen L, Joris G, Cuijt I, Corsmit E, Elinck E, Van Dongen J, Vermeulen S, Van den Broeck M, Vaerenberg C, Mattheijssens M, Peeters K, Robberecht W, Cras P, Martin JJ, De Deyn PP, Cruts M, Van Broeckhoven C. A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study. Lancet Neurol. 2012; 11:54–65.10.1016/S1474-4422(11)70261-7 [PubMed: 22154785]
16. Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, Ogar JM, Rohrer JD, Black S, Boeve BF, Manes F, Dronkers NF, Vandenberghe R, Rascovsky K, Patterson K, Miller BL, Knopman DS, Hodges JR, Mesulam MM, Grossman M. Classification of primary progressive aphasia and its variants. Neurology. 2011; 76:1006–1014.10.1212/WNL.0b013e31821103e6 [PubMed: 21325651]
17. Hutton M, Lendon CL, Rizzu P, Baker M, Froelich S, Houlden H, Pickering-Brown S, Chakraverty S, Isaacs A, Grover A, Hackett J, Adamson J, Lincoln S, Dickson D, Davies P, Petersen RC, Stevens M, de Graaff E, Wauters E, van Baren J, Hillebrand M, Joosse M, Kwon JM, Nowotny P, Che LK, Norton J, Morris JC, Reed LA, Trojanowski J, Basun H, Lannfelt L, Neystat M, Fahn S, Dark F, Tannenberg T, Dodd PR, Hayward N, Kwok JB, Schofield PR, Andreadis A, Snowden J, Craufurd D, Neary D, Owen F, Oostra BA, Hardy J, Goate A, van Swieten J, Mann D, Lynch T, Heutink P. Association of missense and 5′-splice-site mutations in tau with the inherited dementia FTDP-17. Nature. 1998; 393:702–705.10.1038/31508 [PubMed: 9641683]
18. Kern AD, Kondrashov FA. Mechanisms and convergence of compensatory evolution in mammalian mitochondrial tRNAs. Nat Genet. 2004; 36:1207–1212.10.1038/ng1451 [PubMed: 15502829]
19. Kondrashov AS, Sunyaev S, Kondrashov FA. Dobzhansky-Muller incompatibilities in protein evolution. Proc Natl Acad Sci U S A. 2002; 99:14878–14883.10.1073/pnas.232565499 [PubMed: 12403824]
20. Lang CM, Fellerer K, Schwenk BM, Kuhn PH, Kremmer E, Edbauer D, Capell A, Haass C. Membrane orientation and subcellular localization of transmembrane protein 106B (TMEM106B), a major risk factor for frontotemporal lobar degeneration. J Biol Chem. 2012; 287:19355–19365.10.1074/jbc.M112.365098 [PubMed: 22511793]
21. Levine TP, Daniels RD, Gatta AT, Wong LH, Hayes MJ. The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. Bioinformatics. 2013; 29:499–503.10.1093/bioinformatics/bts725 [PubMed: 23329412]

22. Litvan I, Agid Y, Calne D, Campbell G, Dubois B, Duvoisin RC, Goetz CG, Golbe LI, Grafman J, Growdon JH, Hallett M, Jankovic J, Quinn NP, Tolosa E, Zee DS. Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): report of the NINDS-SPSP international workshop. Neurology. 1996; 47:1–9. [PubMed: 8710059]
23. Mackenzie IR, Neumann M, Baborie A, Sampathu DM, Du Plessis D, Jaros E, Perry RH, Trojanowski JQ, Mann DM, Lee VM. A harmonized classification system for FTLD-TDP pathology. Acta Neuropathol. 2011; 122:111–113.10.1007/s00401-011-0845-8 [PubMed: 21644037]
24. McKhann GM, Albert MS, Grossman M, Miller B, Dickson D, Trojanowski JQ. Work Group on Frontotemporal Dementia and Pick's Disease . Clinical and pathological diagnosis of frontotemporal dementia: report of the Work Group on Frontotemporal Dementia and Pick's Disease. Arch Neurol. 2001; 58:1803–1809. [PubMed: 11708987]
25. Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E, Schmid B, Kretzschmar HA, Cruts M, Van Broeckhoven C, Haass C, Edbauer D. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science. 2013; 339:1335–1338.10.1126/science.1232927 [PubMed: 23393093]
26. Nicholson AM, Finch NA, Wojtas A, Baker MC, Perkerson RB, Castanedes-Casey M, Rousseau L, Benussi L, Binetti G, Ghidoni R, Hsiung GY, Mackenzie IR, Finger E, Boeve BF, Ertekin-Taner N, Graff-Radford NR, Dickson DW, Rademakers R. TMEM106B p.T185S regulates TMEM106B protein levels: implications for frontotemporal dementia. J Neurochem. 201310.1111/jnc.12329
27. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, van Swieten JC, Seelaar H, Dopper EG, Onyike CU, Hillis AE, Josephs KA, Boeve BF, Kertesz A, Seeley WW, Rankin KP, Johnson JK, Gorno-Tempini ML, Rosen H, Prioleau-Latham CE, Lee A, Kipps CM, Lillo P, Piguet O, Rohrer JD, Rossor MN, Warren JD, Fox NC, Galasko D, Salmon DP, Black SE, Mesulam M, Weintraub S, Dickerson BC, Diehl-Schmid J, Pasquier F, Deramecourt V, Lebert F, Pijnenburg Y, Chow TW, Manes F, Grafman J, Cappa SF, Freedman M, Grossman M, Miller BL. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain. 2011; 134:2456–2477.10.1093/brain/awr179 [PubMed: 21810890]
28. Renton AE, Majounie E, Waite A, Simon-Sanchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L, Kalimo H, Paetau A, Abramzon Y, Remes AM, Kaganovich A, Scholz SW, Duckworth J, Ding J, Harmer DW, Hernandez DG, Johnson JO, Mok K, Ryten M, Trabzuni D, Guerreiro RJ, Orrell RW, Neal J, Murray A, Pearson J, Jansen IE, Sondervan D, Seelaar H, Blake D, Young K, Halliwell N, Callister JB, Toulson G, Richardson A, Gerhard A, Snowden J, Mann D, Neary D, Nalls MA, Peuralinna T, Jansson L, Isoviita VM, Kaivorinne AL, Holtta-Vuori M, Ikonen E, Sulkava R, Benatar M, Wuu J, Chio A, Restagno G, Borghero G, Sabatelli M, Heckerman D, Rogaeva E, Zinman L, Rothstein JD, Sendtner M, Drepper C, Eichler EE, Alkan C, Abdullaev Z, Pack SD, Dutra A, Pak E, Hardy J, Singleton A, Williams NM, Heutink P, Pickering-Brown S, Morris HR, Tienari PJ, Traynor BJ. ITALSGEN Consortium. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron. 2011; 72:257–268.10.1016/j.neuron.2011.09.010 [PubMed: 21944779]
29. Schenk MF, Szendro IG, Salverda ML, Krug J, de Visser JA. Patterns of Epistasis between Beneficial Mutations in an Antibiotic Resistance Gene. Mol Biol Evol. 201310.1093/molbev/mst096
30. Seelaar H, Rohrer JD, Pijnenburg YA, Fox NC, van Swieten JC. Clinical, genetic and pathological heterogeneity of frontotemporal dementia: a review. J Neurol Neurosurg Psychiatry. 2011; 82:476–486.10.1136/jnnp.2010.212225 [PubMed: 20971753]
31. Shankaran SS, Capell A, Hruscha AT, Fellerer K, Neumann M, Schmid B, Haass C. Missense mutations in the progranulin gene linked to frontotemporal lobar degeneration with ubiquitin-immunoreactive inclusions reduce progranulin production and secretion. J Biol Chem. 2008; 283:1744–1753.10.1074/jbc.M705115200 [PubMed: 17984093]
32. Silva RF, Mendonca SC, Carvalho LM, Reis AM, Gordo I, Trindade S, Dionisio F. Pervasive sign epistasis between conjugative plasmids and drug-resistance chromosomal mutations. PLoS Genet. 2011; 7:e1002181.10.1371/journal.pgen.1002181 [PubMed: 21829372]

33. Skibinski G, Parkinson NJ, Brown JM, Chakrabarti L, Lloyd SL, Hummerich H, Nielsen JE, Hodges JR, Spillantini MG, Thusgaard T, Brandner S, Brun A, Rossor MN, Gade A, Johannsen P, Sorensen SA, Gydesen S, Fisher EM, Collinge J. Mutations in the endosomal ESCRTIII-complex subunit CHMP2B in frontotemporal dementia. Nat Genet. 2005; 37:806–808.10.1038/ng1609 [PubMed: 16041373]
34. Snowden JS, Rollinson S, Thompson JC, Harris JM, Stopford CL, Richardson AM, Jones M, Gerhard A, Davidson YS, Robinson A, Gibbons L, Hu Q, DuPlessis D, Neary D, Mann DM, Pickering-Brown SM. Distinct clinical and pathological characteristics of frontotemporal dementia associated with C9ORF72 mutations. Brain. 2012; 135:693–708.10.1093/brain/awr355 [PubMed: 22300873]
35. Stewart H, Rutherford NJ, Briemberg H, Krieger C, Cashman N, Fabros M, Baker M, Fok A, DeJesus-Hernandez M, Eisen A, Rademakers R, Mackenzie IR. Clinical and pathological features of amyotrophic lateral sclerosis caused by mutation in the C9ORF72 gene on chromosome 9p. Acta Neuropathol. 2012; 123:409–417.10.1007/s00401-011-0937-5 [PubMed: 22228244]
36. Strong MJ, Grace GM, Freedman M, Lomen-Hoerth C, Woolley S, Goldstein LH, Murphy J, Shoesmith C, Rosenfeld J, Leigh PN, Bruijn L, Ince P, Figlewicz D. Consensus criteria for the diagnosis of frontotemporal cognitive and behavioural syndromes in amyotrophic lateral sclerosis. Amyotroph Lateral Scler. 2009; 10:131–146. [PubMed: 19462523]
37. Toledo JB, Van Deerlin VM, Lee EB, Suh E, Baek Y, Robinson JL, Xie SX, McBride J, Wood EM, Schuck T, Irwin DJ, Gross RG, Hurtig H, McCluskey L, Elman L, Karlawish J, Schellenberg G, Chen-Plotkin A, Wolk D, Grossman M, Arnold SE, Shaw LM, Lee VM, Trojanowski JQ. A platform for discovery: The University of Pennsylvania Integrated Neurodegenerative Disease Biobank. Alzheimers Dement. 201310.1016/j.jalz.2013.06.003
38. Van Deerlin VM, Sleiman PM, Martinez-Lage M, Chen-Plotkin A, Wang LS, Graff-Radford NR, Dickson DW, Rademakers R, Boeve BF, Grossman M, Arnold SE, Mann DM, Pickering-Brown SM, Seelaar H, Heutink P, van Swieten JC, Murrell JR, Ghetti B, Spina S, Grafman J, Hodges J, Spillantini MG, Gilman S, Lieberman AP, Kaye JA, Woltjer RL, Bigio EH, Mesulam M, Al-Sarraj S, Troakes C, Rosenberg RN, White CL 3rd, Ferrer I, Llado A, Neumann M, Kretzschmar HA, Hulette CM, Welsh-Bohmer KA, Miller BL, Alzualde A, Lopez de Munain A, McKee AC, Gearing M, Levey AI, Lah JJ, Hardy J, Rohrer JD, Lashley T, Mackenzie IR, Feldman HH, Hamilton RL, Dekosky ST, van der Zee J, Kumar-Singh S, Van Broeckhoven C, Mayeux R, Vonsattel JP, Troncoso JC, Kril JJ, Kwok JB, Halliday GM, Bird TD, Ince PG, Shaw PJ, Cairns NJ, Morris JC, McLean CA, DeCarli C, Ellis WG, Freeman SH, Frosch MP, Growdon JH, Perl DP, Sano M, Bennett DA, Schneider JA, Beach TG, Reiman EM, Woodruff BK, Cummings J, Vinters HV, Miller CA, Chui HC, Alafuzoff I, Hartikainen P, Seilhean D, Galasko D, Masliah E, Cotman CW, Tunon MT, Martinez MC, Munoz DG, Carroll SL, Marson D, Riederer PF, Bogdanovic N, Schellenberg GD, Hakonarson H, Trojanowski JQ, Lee VM. Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. Nat Genet. 2010; 42:234–239.10.1038/ng.536 [PubMed: 20154673]
39. van der Zee J, Van Langenhove T, Kleinberger G, Sleegers K, Engelborghs S, Vandenberghe R, Santens P, Van den Broeck M, Joris G, Brys J, Mattheijssens M, Peeters K, Cras P, De Deyn PP, Cruts M, Van Broeckhoven C. TMEM106B is associated with frontotemporal lobar degeneration in a clinically diagnosed patient cohort. Brain. 2011; 134:808–815.10.1093/brain/awr007 [PubMed: 21354975]
40. Vass R, Ashbridge E, Geser F, Hu WT, Grossman M, Clay-Falcone D, Elman L, McCluskey L, Lee VM, Van Deerlin VM, Trojanowski JQ, Chen-Plotkin AS. Risk genotypes at TMEM106B are associated with cognitive impairment in amyotrophic lateral sclerosis. Acta Neuropathol. 2011; 121:373–380.10.1007/s00401-010-0782-y [PubMed: 21104415]
41. Watts GD, Wymer J, Kovach MJ, Mehta SG, Mumm S, Darvish D, Pestronk A, Whyte MP, Kimonis VE. Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. Nat Genet. 2004; 36:377–381.10.1038/ng1332 [PubMed: 15034582]
42. Weinreich DM, Watson RA, Chao L. Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. Evolution. 2005; 59:1165–1174. [PubMed: 16050094]

43. Wijsman EM, Daw EW, Yu X, Steinbart EJ, Nochlin D, Bird TD, Schellenberg GD. APOE and other loci affect age-at-onset in Alzheimer's disease families with PS2 mutation. Am J Med Genet B Neuropsychiatr Genet. 2005; 132B:14–20.10.1002/ajmg.b.30087 [PubMed: 15389756]

44. Xie SX, Baek Y, Grossman M, Arnold SE, Karlawish J, Siderowf A, Hurtig H, Elman L, McCluskey L, Van Deerlin V, Lee VM, Trojanowski JQ. Building an integrated neurodegenerative disease database at an academic health center. Alzheimers Dement. 2011; 7:e84–93.10.1016/j.jalz.2010.08.233 [PubMed: 21784346]

45. Zhang W, Dourado DF, Fernandes PA, Ramos MJ, Mannervik B. Multidimensional epistasis and fitness landscapes in enzyme evolution. Biochem J. 2012; 445:39–46.10.1042/BJ20120136 [PubMed: 22533640]

INTERNATIONAL COLLABORATION FOR FRONTOTEMPORAL LOBAR DEGENERATION

The International Collaboration for Frontotemporal Lobar Degeneration consisted of clinical sites collaborating to collect cases for an FTLD-TDP genomewide association study (GWAS); this GWAS led to the discovery that common variants in *TMEM106B* are a genetic risk factor for FTLD-TDP [38]. Members of the Collaboration who contributed *C9orf72(+)*FTLD-TDP cases for the current study include Irina Alafuzoff, Anna Antonell, Nenad Bogdanovic, William Brooks, Nigel Cairns, Johnathan Cooper-Knock, Carl W. Cotman, Patrick Cras, Marc Cruts, Peter P. De Deyn, Charles DeCarli, Carol Dobson-Stone, Sebastiaan Engelborghs, Nick Fox, Douglas Galasko, Marla Gearing, Ilse Gijselinck, Jordan Grafman, Paivi Hartikainen, Kimmo J. Hatanpaa, J. Robin Highley, John Hodges, Christine Hulette, Paul G. Ince, Lee-Way Jin, Janine Kirby, Julia Kofler, Jillian Kril, John J. B. Kwok, Allan Levey, Andrew Lieberman, Albert Llado, Jean-Jacques Martin, Eliezer Masliah, Christopher J. McDermott, Catriona McLean, Ann C. McKee, Simon Mead, Carol A. Miller, Josh Miller, David Munoz, Jill Murrell, Henry Paulson, Olivier Piguet, Martin Rossor, Raquel Sanchez-Valle, Mary Sano, Julie Schneider, Lisa Silbert, Salvatore Spina, Julie van der Zee, Tim Van Langenhove, Jason Warren, Stephen B. Wharton, Charles L. White III, Randall Woltjer.

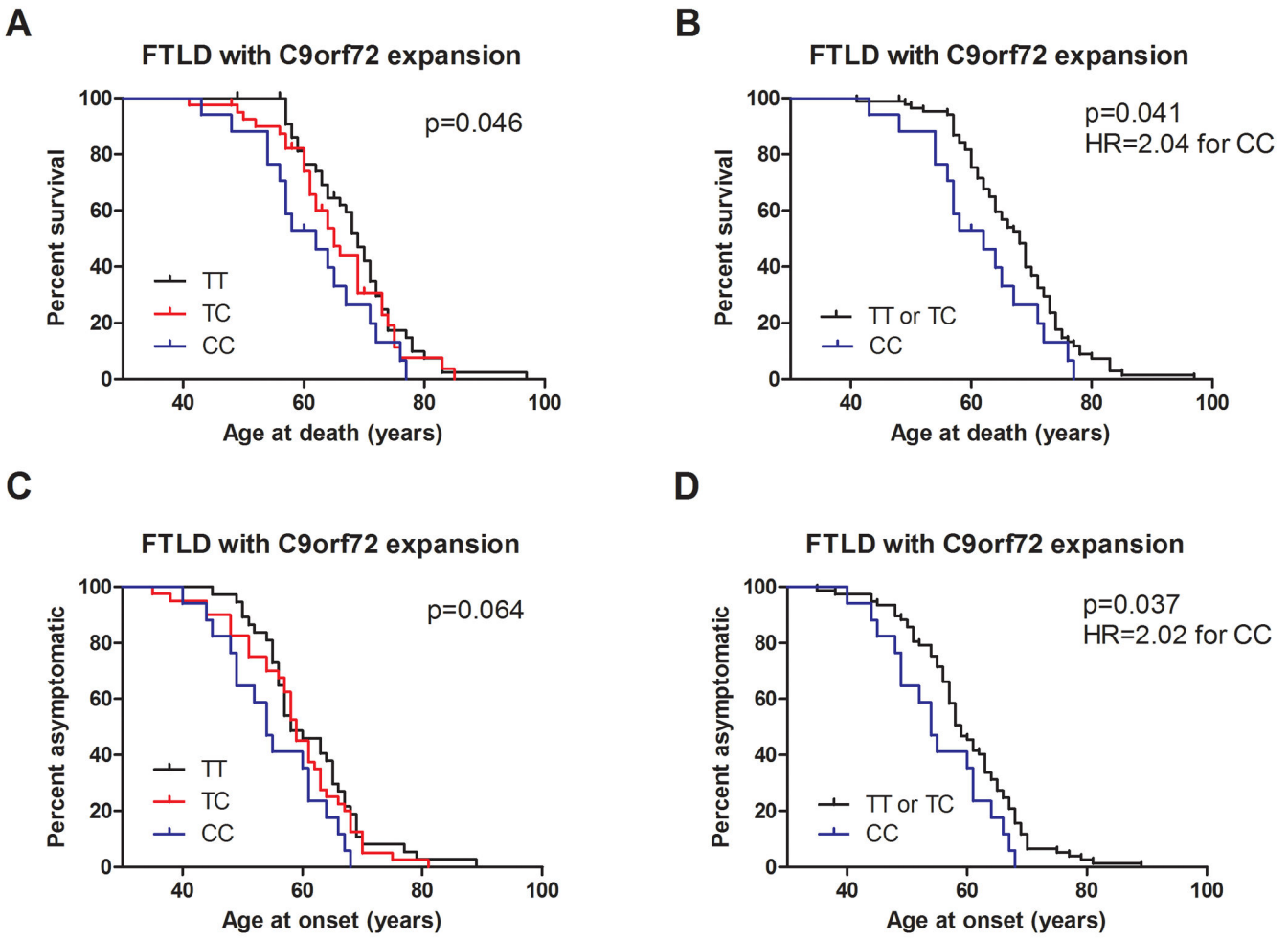


Fig. 1. *TMEM106B* genotype influences age at death and age at onset in *C9orf72(+)* FTLD

All survival analyses were performed in 104 total *C9orf72(+)* FTLD cases, from the combined discovery and replication cohorts. Of these 104 total cases, 89 had available age-at-death data, and 94 had age-at-onset data.

A) Age at death was significantly associated with *TMEM106B* genotype at rs1990622, the top SNP associated with FTLD-TDP in our prior GWAS. Log rank test for trend two-tailed p=0.046, assuming a codominant model.

B) Under a major-allele-dominant model, *TMEM106B* rs1990622 genotype was even more significantly associated with age at death, with more than twice the risk of death at any given age for CC carriers compared to carriers of one or more T alleles (two-tailed p=0.041, HR=2.039, 95% CI 1.031–4.033).

C) Age at onset showed a trend towards association with *TMEM106B* genotype at rs1990622. Log rank test for trend two-tailed p=0.064, assuming a codominant model.

D) Under a major-allele-dominant model, *TMEM106B* rs1990622 genotype showed a significant association with age at disease onset, with more than twice the risk of disease onset at any given age for CC carriers compared to carriers of one or more T alleles (two-tailed p=0.037, HR=2.022, 95% CI 1.042–3.925)

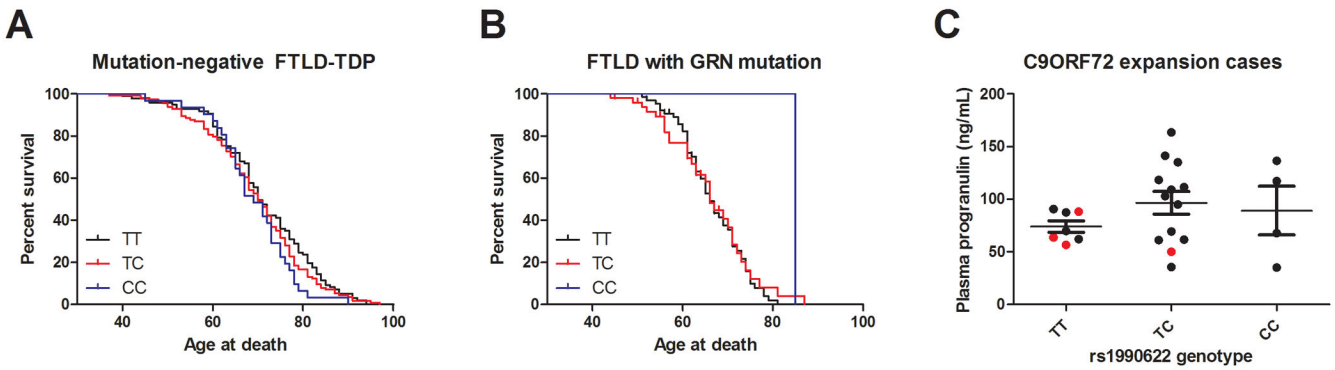


Fig. 2. *TMEM106B* genotype does not affect age at death or age at onset for FTLD-TDP without *C9orf72* expansions

A) In 241 FTLD-TDP cases negative for *GRN* mutations or *C9orf72* expansions, *TMEM106B* genotype at rs1990622 did not affect age at death.

B) In 116 FTLD-TDP cases with *GRN* mutations, we found no significant difference in age at death comparing TT and TC carriers at rs1990622. In this cohort, only one individual had the CC genotype, precluding our ability to evaluate the influence of this genotype.

C) Plasma progranulin levels were measured in a convenience subset of 24 *C9orf72* expansion carriers by ELISA. Progranulin levels did not differ significantly by *TMEM106B* rs1990622 genotype, although the TT carriers exhibited significantly less variance in their progranulin levels. Black dots indicate individuals who presented with ALS, while red dots indicate individuals who presented with FTLD.

Table 1

***TMEM106B* genotype affects age at death in *C9orf72* expansion carriers with FTLD or FTLD-TDP in a discovery cohort**

Linear regressions were used to evaluate the effect of *TMEM106B* genotype at rs1990622 on the age at death or age at onset in *C9orf72* expansion carriers from a discovery cohort. In individuals who presented with clinical FTLD or FTLD-TDP, rs1990622 genotype was significantly associated with age at death in both univariate models and models adjusting for age and presence/absence of motor neuron disease (MND). In individuals who presented with ALS, rs1990622 genotype was not significantly associated with age at death, with a trend towards association with age at onset. Asterisks denote significance.

Disease	Outcome	Predictors	Beta (rs1990622, each major allele)	R^2 for model	P-value (rs1990622)
FTLD and FTLD-TDP	Age at Death (n=14)	rs1990622	+6.278	0.303	**0.024 ***
		rs1990622, Sex, MND	+5.297	0.393	**0.049 ***
	Age at Onset (n=26)	rs1990622		n.s.	
		rs1990622, Sex, MND		n.s.	
ALS	Age at Death (n=39)	rs1990622		n.s.	
		rs1990622, Sex, FTD		n.s.	
	Age at Onset (n=47)	rs1990622	−4.264	0.044	0.085 n.s.
		rs1990622, Sex, FTD	−4.900	0.075	**0.048 ***

Acta Neuropathol. Author manuscript; available in PMC 2014 April 29.

Table 2

***TMEM106B* genotype affects age at death and age at onset in *C9orf72* expansion carriers in a multi-site FTLD-TDP replication cohort**

Linear regressions were used to evaluate the effect of *TMEM106B* genotype at rs1990622 on the age at death or age at onset in *C9orf72(+)* FTLD from a multi-site replication cohort of FTLD-TDP cases. rs1990622 genotype was significantly associated with both age at death and age at onset, in both univariate models and models adjusting for age and presence/absence of motor neuron disease (MND). Asterisks denote significance.

Disease	Outcome	Predictors	Beta (rs1990622, each major allele)	R^2 for model	P-value (rs1990622)
FTLD-TDP	Age at Death (n=75)	rs1990622	+3.342	0.048	**0.016 ***
		rs1990622, Sex, MND	+3.413	0.032	**0.019 ***
	Age at Onset (n=68)	rs1990622	+3.473	0.049	**0.019 ***
		rs1990622, Sex, MND	+3.198	0.057	**0.032 ***

Table 3

***TMEM106B* rs1990622 genotype is associated with FTLD-TDP in all genetic subgroups**

Chi-square tests were performed to evaluate for association between disease and rs1990622 genotype for FTLD-TDP subgroups defined by the presence of *GRN* mutations (*GRN(+)* FTLD-TDP), presence of *C9orf72* expansions (*C9orf72(+)* FTLD-TDP), or the absence of both genetic mutations (FTLD-TDP (no mutation)). The major allele was significantly associated with disease in all three subgroups. Allele frequencies for normal controls provided here are from our previously published GWAS.

Disease status	N	rs1990622 Major allele T	rs1990622 Minor allele C	p-value	Odds ratio	95% CI
Normal	2509	0.564	0.436	-		
***GRN(+)* FTLD-TDP**	116	0.776	0.224	**<0.0001**	**2.675**	**1.955–3.660**
***C9orf72(+)*FTLD-TDP**	80	0.669	0.331	**0.008**	**1.560**	**1.117–2.179**
FTLD-TDP (no mutation)	241	0.640	0.360	**0.001**	**1.375**	**1.131–1.671**

Acta Neuropathol. Author manuscript; available in PMC 2014 April 29.

Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 2014; Early Online: 1–3

informa
healthcare

REPORT

Intermediate length C9orf72 expansion in an ALS patient without classical C9orf72 neuropathology

ALEXANDER M. BEER[1], JOHNATHAN COOPER-KNOCK[1], ADRIAN HIGGINBOTTOM[1], J. ROBIN HIGHLEY[1], STEPHEN B. WHARTON[1], PAUL G. INCE[1], ANTONIO MILANO[2], ASHLEY A. JONES[3], AMMAR AL-CHALABI[3], JANINE KIRBY[1] & PAMELA J. SHAW[1]

[1]*Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, UK,* [2]*Sheffield Diagnostic Genetic Service, Sheffield Children's NHS Foundation Trust, UK, and* [3]*Institute of Psychiatry, King's College London, UK*

Introduction

An intronic GGGGCC hexanucleotide repeat expansion in *C9orf72* is the most common genetic variant of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). Expansions occur in the context of a risk haplotype at locus 9p21 (1). Early work suggested a cut-off of 30 repeats to identify a pathogenic C9orf72 expansion. Most controls have ≤ 3 repeats, but control individuals have been identified with ≥ 30 repeats (2). Understanding the relationship between repeat length and pathogenicity is crucial for translational research.

The 9p21 risk haplotype is associated with C9orf72 repeat expansion lengths greater than 7 units; subsequently 7–24 repeats were defined as intermediate length (3). However, it is unclear whether expansions of this length are pathogenic. There is no significant difference between the proportion of patients and controls carrying intermediate length expansions (3). Reported patients with intermediate length expansions clinically resemble patients with larger expansions including, in certain cases, described family history of disease (4,5). However, given the wide phenotypic spectrum of both C9orf72-ALS and non-C9orf72-ALS, this is not conclusive. While characteristic neuropathology is associated with C9orf72 disease (6,7), until now pathological analysis has been missing from the characterization of patients with intermediate length expansions.

Case report

We report a 75-year-old female who presented with an 18-month history of progressive distal weakness. There was no family history of neuromuscular disease. Examination revealed weakness particularly in the lower limbs, hyperreflexia, wasting and fasciculation. There was no sensory or cognitive impairment. A diagnosis of ALS was made following full neurological investigation. She died from respiratory failure 31 months after symptom onset.

Pathological material was obtained from the Sheffield Brain Tissue Bank. Ethics committee approval and written consent was obtained. Genomic DNA was extracted from cerebellar material and a C9orf72 expansion of 16 GGGGCC repeats was identified by PCR analysis (Figure 1A) and Southern hybridization (8) (Figure 1B). Single nucleotide polymorphisms (SNPs) at rs3849942 and rs2814707 are highly associated with the 9p21 risk allele (1). Genotyping showed that our patient was heterozygous for the minor, or risk, allele at both positions. This is highly suggestive that her intermediate length expansion had occurred on the background of the 9p21 risk haplotype. Taqman allelic discrimination assay on an ABI 7900HT Real-Time PCR system was used for genotyping SNPs. Pre-designed primers and probes were purchased from Applied Biosystems (Foster City, USA).

Histology revealed loss of lower motor neurons in the medulla and spinal cord with Bunina bodies in residual neurons. The motor cortex had detectable superficial cortical vacuolation. Immunohistochemistry revealed p62 and TDP-43-positive neuronal and glial cytoplasmic inclusions in motor cortex, brainstem and spinal cord. C9orf72 disease is associated with an abundance of 'star-shaped' ubiquitylated, TDP-43 negative inclusions in extramotor areas (6);

Correspondence: P. Shaw, Sheffield Institute for Translational Neuroscience (SITraN), Department of Neuroscience, University of Sheffield, 385a Glossop Road, Sheffield S10 2HQ, UK. Fax: +44 0 114 2222290. E-mail: pamela.shaw@sheffield.ac.uk

(Received 15 July 2014; accepted 12 October 2014)

ISSN 2167-8421 print/ISSN 2167-9223 online © 2014 Informa Healthcare
DOI: 10.3109/21678421.2014.977296

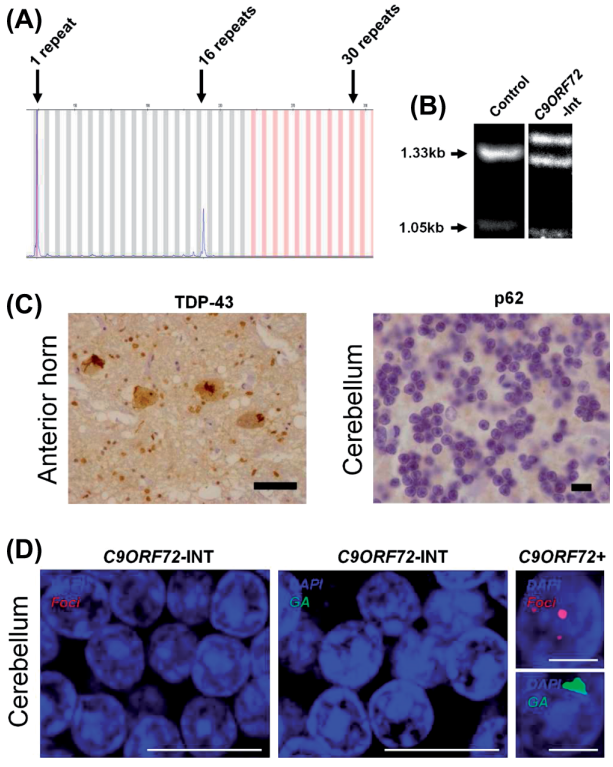


Figure 1. Patient with an intermediate size C9orf72 GGGGCC repeat expansion does not have typical C9orf72 pathology. (A) Genotyping PCR of a wild-type control and our patient. Shaded lines represent numbers of repeats from 1 to 30. The patient has alleles with 1 and 16 repeats. (B) Southern hybridization based detection of the intermediate length C9orf72 allele. Bands at approximately 1.33 kb correspond to EcoRI/XbaI fragments derived from the locus normally containing the C9orf72 expansion. The control case has a single band at 1.33 kb indicating that both alleles are of normal length, while our case with an intermediate length expansion has two bands corresponding to the two alleles of length 1 and 16 repeats. The 1.0 5 kb band is an internal control. (C) Immunohistochemistry revealed classical TDP-43 proteinopathy in motor areas including spinal cord (left panel, scale bar 50 μm), brainstem and motor cortex, but no ubiquitylated, TDP-43-negative inclusions (the hallmark C9orf72 ALS pathology) in extramotor areas including cerebellum (right panel, scale bar 10 μm), hippocampus, substantia nigra and neocortex. (D) Immunohistochemistry of cerebellar granule neurons failed to demonstrate either GGGGCC repeat RNA foci (left panel), or Poly-(Gly-Ala)-dipeptide repeat protein (middle panel), despite examination of >100 cerebellar granule neurons; scale bar 10 μm. Positive staining from a C9orf72-ALS case with >2000 GGGGCC repeats is included for comparison (right panels, scale bar 3 μm).

and sense RNA foci and dipeptide repeat protein within neuronal inclusions in cerebellar granule neurons (7). All of these features were absent in our case (Figure 1C, D). Pathology was examined as described previously (6,9).

Discussion

We have identified a patient with an intermediate length C9orf72 expansion of 16 GGGGCC repeats in association with the 9p21 risk haplotype, who presented with clinically typical ALS. We confirmed the repeat length in the CNS by PCR and Southern hybridization. Moreover, we have demonstrated that this patient did not have the pathology characteristic of C9orf72-ALS.

Our findings suggest that the disease process in our patient was distinct from typical C9orf72-ALS. We propose that this patient suffered sporadic ALS and therefore the intermediate length C9orf72 expansion was not key to pathogenesis. This is consistent with growing evidence for a gain-of-function toxicity mediated either by RNA foci transcribed from the repeat sequence or dipeptide repeat protein formed by repeat associated non-ATG translation (2). In a gain-of-function mechanism the expansion size would be expected to modify the disease phenotype. Neither the case described in this report nor the others presented in the literature (4,5) show reduced clinical severity compared to other C9orf72-disease patients despite >100-fold less of the GGGGCC repeat expansion. The 9p21 risk haplotype may be associated with a propensity for the C9orf72 locus to expand, but our case suggests that this must reach a threshold value, above intermediate repeat length, to initiate typical C9orf72 disease. It should be noted that this is a single case and will require validation. An alternative potential explanation for our findings is that the case described may be an example of pathological heterogeneity in C9orf72 disease. However, in our experience and that of others (7,10), the presence of cytoplasmic Poly-(Gly-Ala)-dipeptide repeat protein and nuclear RNA foci in cerebellar granule neurons, in patients with C9orf72 expansions of >30 repeats, is universal.

Acknowledgements

We acknowledge EU Framework 7 (Euro-motor), JPND/MRC SOPHIA, STRENGTH and ALS-CarE project grants. JCK and JRH hold MNDA/MRC Lady Edith Wolfson Fellowships. We are grateful to patients and controls who donated biosamples for research.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

1. Shatunov A, Mok K, Newhouse S, Weale ME, Smith B, Vance C, et al. Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. Lancet Neurol. 2010; 9:986–94.
2. Cooper-Knock J, Shaw PJ, Kirby J. The widening spectrum of C9orf72 related disease; genotype/phenotype correlations and potential modifiers of clinical phenotype. Acta Neuropathologica. 2014;127:333–45.
3. van der Zee J, Gijselinck I, Dillen L, van Langenhove T, Theuns J, Engelborghs S, et al. A Pan-European Study of the C9orf72 Repeat Associated with FTLD: Geographic Prevalence, Genomic Instability, and Intermediate Repeats. Human Mutation. 2013;34:363–73.

4. Gomez-Tortosa E, Gallego J, Guerrero-Lopez R, Marcos A, Gil-Neciga E, Sainz MJ, et al. C9orf72 hexanucleotide expansions of 20–22 repeats are associated with frontotemporal deterioration. Neurology. 2013;80:366–70.
5. Byrne S, Heverin M, Elamin M, Walsh C, Hardiman O. Intermediate repeat expansion length in C9orf72 may be pathological in amyotrophic lateral sclerosis. Amyotroph Lateral Scler Frontotemporal Degener. 2013;135:751–64.
6. Cooper-Knock J, Hewitt C, Highley JR, Brockington A, Milano A, Man S, et al. Clinicopathological features in amyotrophic lateral sclerosis with expansions in C9orf72. Brain. 2012;135:751–64.
7. Mackenzie IR, Frick PM, Neumann M. The neuropathology associated with repeat expansions in the C9orf72 gene. Acta Neuropathol. 2014;127:347–57.
8. Buchman VL, Cooper-Knock J, Connor-Robson N, Higginbottom A, Kirby J, Razinskaya OD, et al. Simultaneous and independent detection of C9orf72 alleles with low and high number of GGGGCC repeats using an optimized protocol of Southern blot hybridization. Molecular Neurodegeneration. 2013;8:12.
9. Cooper-Knock J, Walsh MJ, Higginbottom A, Robin Highley J, Dickman MJ, Edbauer D, et al. Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. Brain. 2014;137:2040–51.
10. Mackenzie IR, Arzberger T, Kremmer E, Troost D, Lorenzl S, Mori K, et al. Dipeptide repeat protein pathology in C9orf72 mutation cases: clinicopathological correlations. Acta Neuropathol. 2013;126:859–79.

5. Concluding discussion – implications for translational work

The discovery of the *C9orf72* genetic variant has opened a new chapter in the search for therapies for both ALS and FTD. Not only are patients with this genetic variant relatively common, accounting for approximately 10% of all ALS patients in our population (**paper 1,** Cooper-Knock *et al.* 2012), but also these patients encompass the clinical and pathological spectrum seen in the more numerous sporadic disease (Cooper-Knock *et al.* 2012). This was not true for the SOD1 genetic variant which has been the mainstay of translational research in ALS for the last 20 years. The suggestion is that *C9orf72*-ALS will share features of pathogenesis with the sporadic disease which means that therapies targeted at *C9orf72* expansions could eventually lead to treatment for almost all ALS patients.

Modifiers of C9orf72-disease severity

Significant questions are raised by both the variability in the C9orf72-disease phenotype and the similarities between *C9orf72*-ALS/FTD and other causes of these diseases. This variability suggests that multiple modifiers might be at work; indeed modifiers may be crucial to disease penetrance, particularly when it is considered that as many as 0.5% of controls may carry an expansion (**paper 1**, Cooper-Knock *et al.* 2012). In **paper 8** and **paper 10** we have investigated the effect of expansion length on various molecular phenotypes and suggest that smaller expansions are not sufficient to reduce transcription of *C9orf72* mRNA (Cooper-Knock *et al.* 2013) or to initiate *C9orf72*-neuropathology (Beer *et al.* 2014). In the search for therapeutics it is important to know if expansion length is a key determinant of disease; our work suggests that longer repeat lengths are necessary to produce true *C9orf72*-disease. The identification of SNPs in TMEM106B as a modifier of *C9orf72*-FTD in **paper 11** (Gallagher *et al.* 2014), is a first step towards a novel therapy; the fact that SNPs in

this gene are not a modifier of *C9orf72*-ALS is not yet understood. The identification of modifiers is likely to be crucial in design of disease models and potential therapeutics.

Variability in the clinical phenotype of C9orf72-disease

If *C9orf72*-disease is to become penetrant, what determines the particular phenotype? Why do *C9orf72*-patients, even within the same family, present with significantly different symptoms including ALS, FTD, Huntington's phenocopy and neuropsychiatric disease. No conclusive differences have been demonstrated in the size or nature of the expansion according to phenotype, even between the most common ALS and FTD presentations (Dols-Icardo *et al.* 2013, van Blitterswijk *et al.* 2013). The 'prionoid' hypothesis suggests that disease progression in both ALS and FTD might be governed by self-propagation of misfolded protein along predefined anatomical pathways (Jucker *et al.* 2013). If this applies to *C9orf72*-disease then the key difference between phenotypes may be in the site of initiation. It is conceivable that *C9orf72* mutation applies a global insult to all areas of the CNS but disease initiation occurs only once a specific area crosses some threshold determined by a combination of selective (and individual) vulnerability. Moreover, other synergistic insults may have a role, such as other genetic mutations (van Blitterswijk *et al.* 2012). Populations of motor neurons and cortical neurons may both be more vulnerable to the *C9orf72*-insult than for example sensory neurons, but which is most vulnerable i.e. the site of disease initiation, may vary according to a combination of an individual's genetic background and their specific environmental exposures. In this respect it is tempting to hypothesise which environmental factors might have a role, for example exercise may predispose an individual to develop ALS (Harwood *et al.* 2009) whereas cognitive reserve may modify the risk of FTD (Premi *et al.* 2013).

In modelling of *C9orf72*-disease perhaps further insults, genetic or environmental, will be necessary in addition to expression of the GGGGCC-repeat expansion, to precipitate a particular phenotype.

Commonality between C9orf72-disease and other genetic variants

What about the similarities between *C9orf72*-disease and both sporadic and other genetic forms of ALS and FTD; what does this tell us about the disease process? A big part of this question involves TDP-43-pathology. Mislocalisation of TDP-43 into neuronal cytoplasmic inclusions is the hallmark of a large proportion of ALS and FTD including *C9orf72*-disease (Neumann *et al.* 2006, Cooper-Knock *et al.* 2012). If TDP-43 pathology is relatively non-specific then this similarity may be of little consequence i.e. if mislocalisation of TDP-43 is not a significant cause of neuronal death, but merely a downstream consequence, then mechanisms of neurotoxicity in the various forms of these diseases may be quite different and different therapeutic approaches may be required for each. This would significantly limit the translational potential of study of *C9orf72*-disease. Indeed there is some evidence for this: TDP-43 pathology is found in other quite different phenotypes including chronic traumatic encephalopathy (McKee *et al.* 2015) and mislocalisation of TDP-43 is observed in response to fairly generic cell stress (Colombrita *et al.* 2009). However other significant observations argue for a direct role of TDP-43 in neurotoxicity in ALS/FTD: firstly the similarity between the clinical phenotypes produced by the various forms of the disease. Indeed *C9orf72*-disease is clinically indistinguishable from sporadic disease, which is consistent with a shared mechanism of neurotoxicity. The other important observation is that mutations in *TARDBP*, the gene encoding TDP-43, are themselves a cause of ALS and FTD (Sreedharan *et al.* 2008, Borroni *et al.* 2009); this is compelling evidence that TDP-43 is key to the pathogenesis of

these diseases and therefore that pathogenic mechanisms are shared by *C9orf72*-disease and other forms.

The future of C9orf72-disease research

How do these observations shape the future of *C9orf72*-disease research? The design of disease models is perhaps the most important next step. Much work since the discovery of *C9orf72* expansions has focused on the pathogenic mechanisms. Taking precedence from other neuromuscular diseases caused by repeat expansions, evidence has been gathered for three prominent mechanisms (Cooper-Knock, Kirby *et al.* 2015): RNA toxicity, protein toxicity and haploinsufficiency – in each case parallels can be drawn with mechanisms implicated in ALS more broadly and described in section 2. In **paper 5** and **paper 6** we have characterised the interactions and distribution of sense and antisense RNA foci and of dipeptide repeat proteins (DPRs) derived from each RNA species (Cooper-Knock *et al.* 2014). We have described how both varieties of RNA foci have the potential to sequester RRM-containing proteins important to RNA splicing. Extending this further, in **paper 7**, we have described an increase in the splicing error rate in samples derived from *C9orf72*-ALS patients, which correlates with disease severity. Finally, in **paper 6**, we have shown that motor neurons of the ventral horn express antisense RNA foci and antisense RNA derived DPRs at a higher rate than the sense varieties suggesting that they should be the focus of ongoing translational research in ALS at least. However, whilst worthwhile, ablation of toxic RNA or protein species specific to *C9orf72*-disease is unlikely to lead to treatments which are immediately relevant to the wider disease. Perhaps a better approach would be start with what is shared – TDP-43 pathology – and work to ameliorate that first. Then the key question is about the mechanism of TDP-43 related neurotoxicity: for example, is it loss of TDP-43

from the nucleus or gain of TDP-43 in the cytoplasm? Certain models of *C9orf72*-diesase are described (e.g. Donnelly *et al* 2013, Chew *et al.* 2015) but currently missing is a mammalian model which faithfully recapitulates the whole clinical and pathological phenotype in a similar manner to the G90A-SOD1 mouse model; perhaps *C9orf72*-disease models will be key to answering not only questions about *C9orf72* expansions but also TDP-43 pathology. Very recently a pathological study has highlighted the relative prominence of TDP-43 inclusions in *C9orf72*-ALS, in contrast with dipeptide repeat protein pathology which they conclude is rare and unlikely to be key to neurodegeneration (Gomez-Deza *et al.* 2015).

6. References

Aggarwal S and Cudkowicz M. ALS drug development: reflections from the past and a way forward. Neurotherapeutics 2008; 5: 516-527.

Al-Chalabi A, Fang F, Hanby MF, Leigh PN, Shaw CE, Ye W*, et al.* An estimate of amyotrophic lateral sclerosis heritability using twin data. J Neurol Neurosurg Psychiatry 2010; 81: 1324-1326.

Al-Sarraj S, King A, Troakes C, Smith B, Maekawa S, Bodi I*, et al.* p62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of C9orf72-linked FTLD and MND/ALS. Acta Neuropathol 2011; 122: 691-702.

Alexianu ME, Kozovska M and Appel SH. Immune reactivity in a mouse model of familial ALS correlates with disease progression. Neurology 2001; 57: 1282-1289.

Andersen P. Amyotrophic lateral sclerosis associated with mutations in the CuZn superoxide dismutase gene. Current Neurology and Neuroscience Reports 2006; 6: 37-46.

Appel SH, Beers DR and Henkel JS. T cell-microglial dialogue in Parkinson's disease and amyotrophic lateral sclerosis: are we listening? Trends Immunol 2010; 31: 7-17.

Banerjee R, Mosley RL, Reynolds AD, Dhar A, Jackson-Lewis V, Gordon PH*, et al.* Adaptive immune neuroprotection in G93A-SOD1 amyotrophic lateral sclerosis mice. PLoS One 2008; 3: e2740.

Baralle M, Pastor T, Bussani E and Pagani F. Influence of Friedreich ataxia GAA noncoding repeat expansions on pre-mRNA processing. Am J Hum Genet 2008; 83: 77-88.

Battistini S, Benigni M, Ricci C and Rossi A. SOD1 Mutations in Amyotrophic Lateral Sclerosis. European Neurological Journal 2010;

Beck J, Poulter M, Hensman D, Rohrer JD, Mahoney CJ, Adamson G*, et al.* Large C9orf72 hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population. Am J Hum Genet 2013; 92: 345-353.

Beer AM, Cooper-Knock J, Higginbottom A, Highley JR, Wharton SB, Ince PG*, et al.* Intermediate length C9orf72 expansion in an ALS patient without classical C9orf72 neuropathology. Amyotroph Lateral Scler Frontotemporal Degener 2014; 1-3.

Beers DR, Henkel JS, Zhao W, Wang J and Appel SH. CD4+ T cells support glial neuroprotection, slow disease progression, and modify glial morphology in an animal model of inherited ALS. Proc Natl Acad Sci U S A 2008; 105: 15558-15563.

Belzil VV, Bauer PO, Prudencio M, Gendron TF, Stetler CT, Yan IK*, et al.* Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. Acta Neuropathol 2013; 126: 895-905.

Bensimon G, Lacomblez L and Meininger V. A controlled trial of riluzole in amyotrophic lateral sclerosis. ALS/Riluzole Study Group. N Engl J Med 1994; 330: 585-591.

Benussi L, Rossi G, Glionna M, Tonoli E, Piccoli E, Fostinelli S*, et al.* C9ORF72 Hexanucleotide Repeat Number in Frontotemporal Lobar Degeneration: A Genotype-Phenotype Correlation Study. J Alzheimers Dis 2013;

Blair IP, Williams KL, Warraich ST, Durnall JC, Thoeng AD, Manavis J*, et al.* FUS mutations in amyotrophic lateral sclerosis: clinical, pathological, neurophysiological and genetic analysis. J Neurol Neurosurg Psychiatry 2010; 81: 639-645.

Boeve BF, Boylan KB, Graff-Radford NR, DeJesus-Hernandez M, Knopman DS, Pedraza O*, et al.* Characterization of frontotemporal dementia and/or amyotrophic lateral sclerosis associated with the GGGGCC repeat expansion in C9ORF72. Brain 2012; 135: 765-783.

Borroni B, Bonvicini C, Alberici A, Buratti E, Agosti C, Archetti S, et al. Mutation within TARDBP leads to frontotemporal dementia without motor neuron disease. Hum Mutat 2009; 30: E974-983.

Brady OA, Meng P, Zheng Y, Mao Y and Hu F. Regulation of TDP-43 aggregation by phosphorylation and p62/SQSTM1. J Neurochem 2011; 116: 248-259.

Brettschneider J, Arai K, Del Tredici K, Toledo JB, Robinson JL, Lee EB*, et al.* TDP-43 pathology and neuronal loss in amyotrophic lateral sclerosis spinal cord. Acta Neuropathol 2014; 128: 423-437.

Brettschneider J, Del Tredici K, Toledo JB, Robinson JL, Irwin DJ, Grossman M*, et al.* Stages of pTDP-43 pathology in amyotrophic lateral sclerosis. Ann Neurol 2013; 74: 20-38.

Brettschneider J, Del Tredici K, Toledo JB, Robinson JL, Irwin DJ, Grossman M*, et al.* Stages of pTDP-43 pathology in amyotrophic lateral sclerosis. Ann Neurol 2013;

Bubulya PA, Prasanth KV, Deerinck TJ, Gerlich D, Beaudouin J, Ellisman MH*, et al.* Hypophosphorylated SR splicing factors transiently localize around active nucleolar organizing regions in telophase daughter nuclei. J Cell Biol 2004; 167: 51-63.

Buchman VL, Cooper-Knock J, Connor-Robson N, Higginbottom A, Kirby J, Razinskaya OD*, et al.* Simultaneous and independent detection of C9ORF72 alleles with low and high number of GGGGCC repeats using an optimised protocol of Southern blot hybridisation. Mol Neurodegener 2013; 8: 12.

Buratti E, Brindisi A, Giombi M, Tisminetzky S, Ayala YM and Baralle FE. TDP-43 binds heterogeneous nuclear ribonucleoprotein A/B through its C-terminal tail: an important region for the inhibition of cystic fibrosis transmembrane conductance regulator exon 9 splicing. J Biol Chem 2005; 280: 37572-37584.

Burnett R, Melander C, Puckett JW, Son LS, Wells RD, Dervan PB*, et al.* DNA sequence-specific polyamides alleviate transcription inhibition associated with long GAA.TTC repeats in Friedreich's ataxia. Proc Natl Acad Sci U S A 2006; 103: 11497-11502.

Byrne S, Elamin M, Bede P and Hardiman O. Absence of consensus in diagnostic criteria for familial neurodegenerative diseases. J Neurol Neurosurg Psychiatry 2012; 83: 365-367.

Byrne S, Elamin M, Bede P, Shatunov A, Walsh C, Corr B*, et al.* Cognitive and clinical characteristics of patients with amyotrophic lateral sclerosis carrying a C9orf72 repeat expansion: a population-based cohort study. Lancet Neurol 2012; 11: 232-240.

Byrne S, Heverin M, Elamin M, Walsh C and Hardiman O. Intermediate repeat expansion length in C9orf72 may be pathological in amyotrophic lateral sclerosis. Amyotroph Lateral Scler Frontotemporal Degener 2013;

Cacace R, Van Cauwenberghe C, Bettens K, Gijselinck I, van der Zee J, Engelborghs S*, et al.* C9orf72 G4C2 repeat expansions in Alzheimer's disease and mild cognitive impairment. Neurobiol Aging 2013; 34: 1712 e1711-1717.

Chew J, Gendron TF, Prudencio M, Sasaguri H, Zhang YJ, Castanedes-Casey M, et al. Neurodegeneration. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. Science 2015; 348: 1151-1154.

Chio A, Benzi G, Dossena M, Mutani R and Mora G. Severely increased risk of amyotrophic lateral sclerosis among Italian professional football players. Brain 2005; 128: 472-476.

Chio A, Borghero G, Restagno G, Mora G, Drepper C, Traynor BJ*, et al.* Clinical characteristics of patients with familial amyotrophic lateral sclerosis carrying the pathogenic GGGGCC hexanucleotide repeat expansion of C9ORF72. Brain 2012; 135: 784-793.

Clark RM, De Biase I, Malykhina AP, Al-Mahdawi S, Pook M and Bidichandani SI. The GAA triplet-repeat is unstable in the context of the human FXN locus and displays age-dependent expansions in cerebellum and DRG in a transgenic mouse model. Hum Genet 2007; 120: 633-640.

Cohen TJ, Hwang AW, Restrepo CR, Yuan CX, Trojanowski JQ and Lee VM. An acetylation switch controls TDP-43 function and aggregation propensity. Nat Commun 2015; 6: 5845.

Colombrita C, Zennaro E, Fallini C, Weber M, Sommacal A, Buratti E, et al. TDP-43 is recruited to stress granules in conditions of oxidative insult. J Neurochem 2009; 111: 1051-1061.

Cooper-Knock J, Frolov A, Highley JR, Charlesworth G, Kirby J, Milano A*, et al.* C9ORF72 expansions, parkinsonism, and Parkinson disease: a clinicopathologic study. Neurology 2013; 81: 808-811.

Cooper-Knock J, Hewitt C, Highley JR, Brockington A, Milano A, Man S*, et al.* Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. Brain 2012; 135: 751-764.

Cooper-Knock J, Higginbottom A, Connor-Robson N, Bayatti N, Bury JJ, Kirby J*, et al.* C9ORF72 transcription in a frontotemporal dementia case with two expanded alleles. Neurology 2013;

Cooper-Knock J, Jenkins T and Shaw PJ. Clinical and Molecular Aspects of Motor Neuron Disease. Colloquium Series on Genomic and Molecular Medicine 2013; 2: 1-60.

Cooper-Knock J, Shaw PJ and Kirby J. The widening spectrum of C9ORF72-related disease; genotype/phenotype correlations and potential modifiers of clinical phenotype. Acta Neuropathol 2014;

Cooper-Knock J, Walsh MJ, Higginbottom A, Highley JR, Dickman MJ, Edbauer D*, et al.* Sequestration of multiple RNA Recognition Motif-containing proteins by C9ORF72 repeat expansions 2014; Brain 2014; 137:2040-51

Cooper-Knock J, Kirby J, Highley JR and Shaw PJ. The spectrum of C9orf72 mediated neurodegeneration and amyotrophic lateral sclerosis (ALS). Neurotherapeutics. 2015 Mar 3.

Couthouis J, Hart MP, Shorter J, DeJesus-Hernandez M, Erion R, Oristano R*, et al.* A yeast functional screen predicts new candidate ALS disease genes. Proceedings of the National Academy of Sciences 2011; 108: 20881-20890.

Cozzolino M and Carri MT. Mitochondrial dysfunction in ALS. Prog Neurobiol 2012; 97: 54-66.

Da Cruz S and Cleveland DW. Understanding the role of TDP-43 and FUS/TLS in ALS and beyond. Curr Opin Neurobiol 2011; 21: 904-919.

Davidson YS, Barker H, Robinson AC, Thompson JC, Harris J, Troakes C*, et al.* Brain distribution of dipeptide repeat proteins in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. Acta Neuropathol Commun 2014; 2: 70.

Debray S, Race V, Crabbe V, Herdewyn S, Matthijs G, Goris A*, et al.* Frequency of C9orf72 repeat expansions in amyotrophic lateral sclerosis: a Belgian cohort study. Neurobiol Aging 2013; 34: 2890 e2897-2890 e2812.

DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ*, et al.* Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron 2011; 72: 245-256.

del Aguila MA, Longstreth WT, Jr., McGuire V, Koepsell TD and van Belle G. Prognosis in amyotrophic lateral sclerosis: a population-based study. Neurology 2003; 60: 813-819.

Deng H, Gao K and Jankovic J. The role of FUS gene variants in neurodegenerative diseases. Nat Rev Neurol 2014; 10: 337-348.

Deng HX, Chen W, Hong ST, Boycott KM, Gorrie GH, Siddique N*, et al.* Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. Nature 2011; 477: 211-215.

Dobson-Stone C, Hallupp M, Bartley L, Shepherd CE, Halliday GM, Schofield PR*, et al.* C9ORF72 repeat expansion in clinical and neuropathologic frontotemporal dementia cohorts. Neurology 2012; 79: 995-1001.

Dols-Icardo O, Garcia-Redondo A, Rojas-Garcia R, Sanchez-Valle R, Noguera A, Gomez-Tortosa E*, et al.* Characterization of the repeat expansion size in C9orf72 in amyotrophic lateral sclerosis and frontotemporal dementia. Hum Mol Genet 2013;

Donnelly CJ, Zhang PW, Pham JT, Heusler AR, Mistry NA, Vidensky S*, et al.* RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. Neuron 2013; 80: 415-428.

Elden AC, Kim HJ, Hart MP, Chen-Plotkin AS, Johnson BS, Fang X*, et al.* Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. Nature 2010; 466: 1069-1075.

Fahey C, Byrne S, McLaughlin R, Kenna K, Shatunov A, Donohoe G*, et al.* Analysis of the hexanucleotide repeat expansion and founder haplotype at C9ORF72 in an Irish psychosis case-control sample. Neurobiol Aging 2014; 35: 1510 e1511-1515.

Farg MA, Sundaramoorthy V, Sultana JM, Yang S, Atkinson RA, Levina V*, et al.* C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. Hum Mol Genet 2014; 23: 3579-3595.

Ferrante RJ, Browne SE, Shinobu LA, Bowling AC, Baik MJ, MacGarvey U*, et al.* Evidence of increased oxidative damage in both sporadic and familial amyotrophic lateral sclerosis. J Neurochem 1997; 69: 2064-2074.

Ferrante RJ, Shinobu LA, Schulz JB, Matthews RT, Thomas CE, Kowall NW*, et al.* Increased 3-nitrotyrosine and oxidative damage in mice with a human copper/zinc superoxide dismutase mutation. Ann Neurol 1997; 42: 326-334.

Ferrari R, Kero M, Mok K, Paetau A, Tienari PJ, Tynninen O*, et al.* Familial frontotemporal dementia associated with C9orf72 repeat expansion and dysplastic gangliocytoma. Neurobiol Aging 2014; 35: 444 e411-444.

Figlewicz DA, Krizus A, Martinoli MG, Meininger V, Dib M, Rouleau GA*, et al.* Variants of the heavy neurofilament subunit are associated with the development of amyotrophic lateral sclerosis. Hum Mol Genet 1994; 3: 1757-1761.

Fogel BL, Pribadi M, Pi S, Perlman SL, Geschwind DH and Coppola G. C9ORF72 expansion is not a significant cause of sporadic spinocerebellar ataxia. Mov Disord 2012; 27: 1832-1833.

Fratta P, Poulter M, Lashley T, Rohrer JD, Polke JM, Beck J*, et al.* Homozygosity for the C9orf72 GGGGCC repeat expansion in frontotemporal dementia. Acta Neuropathol 2013; 126: 401-409.

Galimberti D, Reif A, Dell'osso B, Kittel-Schneider S, Leonhard C, Herr A*, et al.* The C9ORF72 hexanucleotide repeat expansion is a rare cause of schizophrenia. Neurobiol Aging 2014; 35: 1214 e1217-1214 e1210.

Gallagher MD, Suh E, Grossman M, Elman L, McCluskey L, Van Swieten JC*, et al.* TMEM106B is a genetic modifier of frontotemporal lobar degeneration with C9orf72 hexanucleotide repeat expansions. Acta Neuropathol 2013; (In press):

Gallagher MD, Suh E, Grossman M, Elman L, McCluskey L, Van Swieten JC*, et al.* TMEM106B is a genetic modifier of frontotemporal lobar degeneration with C9orf72 hexanucleotide repeat expansions. Acta Neuropathol 2014;

Garcia-Redondo A, Dols-Icardo O, Rojas-Garcia R, Esteban-Perez J, Cordero-Vazquez P, Munoz-Blanco JL*, et al.* Analysis of the C9orf72 gene in patients with amyotrophic lateral sclerosis in Spain and different populations worldwide. Hum Mutat 2013; 34: 79-82.

Gendron TF, Bieniek KF, Zhang YJ, Jansen-West K, Ash PE, Caulfield T*, et al.* Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. Acta Neuropathol 2013; 126: 829-844.

Goldman JS, Quinzii C, Dunning-Broadbent J, Waters C, Mitsumoto H, Brannagan TH, 3rd*, et al.* Multiple system atrophy and amyotrophic lateral sclerosis in a family with hexanucleotide repeat expansions in C9orf72. JAMA Neurol 2014; 71: 771-774.

Gomez-Tortosa E, Gallego J, Guerrero-Lopez R, Marcos A, Gil-Neciga E, Sainz MJ*, et al.* C9ORF72 hexanucleotide expansions of 20-22 repeats are associated with frontotemporal deterioration. Neurology 2013; 80: 366-370.

Gurney ME, Cutting FB, Zhai P, Andrus PK and Hall ED. Pathogenic mechanisms in familial amyotrophic lateral sclerosis due to mutation of Cu, Zn superoxide dismutase. Pathol Biol (Paris) 1996; 44: 51-56.

Haeusler AR, Donnelly CJ, Periz G, Simko EA, Shaw PG, Kim MS*, et al.* C9orf72 nucleotide repeat structures initiate molecular cascades of disease. Nature 2014; 507: 195-200.

Halliday G, Bigio EH, Cairns NJ, Neumann M, Mackenzie IR and Mann DM. Mechanisms of disease in frontotemporal lobar degeneration: gain of function versus loss of function effects. Acta Neuropathol 2012; 124: 373-382.

Harms M, Benitez BA, Cairns N, Cooper B, Cooper P, Mayo K*, et al.* C9orf72 hexanucleotide repeat expansions in clinical Alzheimer disease. JAMA Neurol 2013; 70: 736-741.

Harwood CA, McDermott CJ and Shaw PJ. Physical activity as an exogenous risk factor in motor neuron disease (MND): a review of the evidence. Amyotroph Lateral Scler 2009; 10: 191-204.

Haverkamp LJ, Appel V and Appel SH. Natural history of amyotrophic lateral sclerosis in a database population. Validation of a scoring system and a model for survival prediction. Brain 1995; 118 (Pt 3): 707-719.

Hensman Moss DJ, Poulter M, Beck J, Hehir J, Polke JM, Campbell T*, et al.* C9orf72 expansions are the most common genetic cause of Huntington disease phenocopies. Neurology 2014; 82: 292-299.

Hortobagyi T, Troakes C, Nishimura AL, Vance C, van Swieten JC, Seelaar H*, et al.* Optineurin inclusions occur in a minority of TDP-43 positive ALS and FTLD-TDP cases and are rarely observed in other neurodegenerative disorders. Acta Neuropathol 2011; 121: 519-527.

Hosler BA, Siddique T, Sapp PC, Sailor W, Huang MC, Hossain A*, et al.* Linkage of familial amyotrophic lateral sclerosis with frontotemporal dementia to chromosome 9q21-q22. Jama 2000; 284: 1664-1669.

Howlett WP, Brubaker GR, Mlingi N and Rosling H. Konzo, an epidemic upper motor neuron disease studied in Tanzania. Brain 1990; 113 (Pt 1): 223-235.

Huey ED, Nagy PL, Rodriguez-Murillo L, Manoochehri M, Goldman J, Lieberman J*, et al.* C9ORF72 repeat expansions not detected in a group of patients with schizophrenia. Neurobiol Aging 2013; 34: 1309 e1309-1310.

Ismail A, Cooper-Knock J, Highley JR, Milano A, Kirby J, Goodall E*, et al.* Concurrence of multiple sclerosis and amyotrophic lateral sclerosis in patients with hexanucleotide repeat expansions of C9ORF72. J Neurol Neurosurg Psychiatry 2013; 84: 79-87.

Jiao B, Guo JF, Wang YQ, Yan XX, Zhou L, Liu XY*, et al.* C9orf72 mutation is rare in Alzheimer's disease, Parkinson's disease, and essential tremor in China. Front Cell Neurosci 2013; 7: 164.

Jucker M and Walker LC. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. Nature 2013; 501: 45-51.

Johnson BS, Snead D, Lee JJ, McCaffery JM, Shorter J and Gitler AD. TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. J Biol Chem 2009; 284: 20329-20339.

Johnson JO, Mandrioli J, Benatar M, Abramzon Y, Van Deerlin VM, Trojanowski JQ*, et al.* Exome sequencing reveals VCP mutations as a cause of familial ALS. Neuron 2010; 68: 857-864.

Johnston CA, Stanton BR, Turner MR, Gray R, Blunt AH, Butt D*, et al.* Amyotrophic lateral sclerosis in an urban setting: a population based study of inner city London. J Neurol 2006; 253: 1642-1643.

Jones AR, Woollacott I, Shatunov A, Cooper-Knock J, Buchman V, Sproviero W*, et al.* Residual association at C9orf72 suggests an alternative amyotrophic lateral sclerosis-causing hexanucleotide repeat. Neurobiol Aging 2013; 34: 2234 e2231-2237.

Kaivorinne AL, Bode MK, Paavola L, Tuominen H, Kallio M, Renton AE*, et al.* Clinical Characteristics of C9ORF72-Linked Frontotemporal Lobar Degeneration. Dement Geriatr Cogn Dis Extra 2013; 3: 251-262.

Katz JS, Katzberg HD, Woolley SC, Marklund SL and Andersen PM. Combined fulminant frontotemporal dementia and amyotrophic lateral sclerosis associated with an I113T SOD1 mutation. Amyotrophic Lateral Sclerosis 2012; 13: 567-569.

Kawamata T, Akiyama H, Yamada T and McGeer PL. Immunologic reactions in amyotrophic lateral sclerosis brain and spinal cord tissue. Am J Pathol 1992; 140: 691-707.

Kim HJ, Kim NC, Wang YD, Scarborough EA, Moore J, Diaz Z*, et al.* Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. Nature 2013; 495: 467-473.

King A, Al-Sarraj S and Shaw C. Frontotemporal lobar degeneration with ubiquitinated tau-negative inclusions and additional alpha-synuclein pathology but also unusual cerebellar ubiquitinated p62-positive, TDP-43-negative inclusions. Neuropathology 2009; 29: 466-471.

King A, Maekawa S, Bodi I, Troakes C and Al-Sarraj S. Ubiquitinated, p62 immunopositive cerebellar cortical neuronal inclusions are evident across the spectrum of TDP-43 proteinopathies but are only rarely additionally immunopositive for phosphorylation-dependent TDP-43. Neuropathology 2011; 31: 239-249.

Kirby J, Goodall E, Smith W, Highley JR, Masanzu R, Hartley J*, et al.* Broad clinical phenotypes associated with TAR-DNA binding protein (TARDBP) mutations in amyotrophic lateral sclerosis. Neurogenetics 2010; 11: 217-225.

Kohli MA, John-Williams K, Rajbhandary R, Naj A, Whitehead P, Hamilton K*, et al.* Repeat expansions in the C9ORF72 gene contribute to Alzheimer's disease in Caucasians. Neurobiol Aging 2013; 34: 1519 e1515-1512.

Konno T, Shiga A, Tsujino A, Sugai A, Kato T, Kanai K*, et al.* Japanese amyotrophic lateral sclerosis patients with GGGGCC hexanucleotide repeat expansion in C9ORF72. J Neurol Neurosurg Psychiatry 2013; 84: 398-401.

Kwon I, Xiang S, Kato M, Wu L, Theodoropoulos P, Wang T*, et al.* Poly-dipeptides encoded by the C9ORF72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. Science 2014;

Lagier-Tourenne C, Baughn M, Rigo F, Sun S, Liu P, Li HR*, et al.* Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. Proc Natl Acad Sci U S A 2013;

Lagier-Tourenne C, Polymenidou M, Hutt KR, Vu AQ, Baughn M, Huelga SC*, et al.* Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. Nat Neurosci 2012; 15: 1488-1497.

Lattante S, Millecamps S, Stevanin G, Rivaud-Pechoux S, Moigneu C, Camuzat A*, et al.* Contribution of ATXN2 intermediary polyQ expansions in a spectrum of neurodegenerative disorders. Neurology 2014;

Lee YB, Chen HJ, Peres JN, Gomez-Deza J, Attig J, Stalekar M*, et al.* Hexanucleotide Repeats in ALS/FTD Form Length-Dependent RNA Foci, Sequester RNA Binding Proteins, and Are Neurotoxic. Cell Rep 2013; 5: 1178-1186.

Lesage S, Le Ber I, Condroyer C, Broussolle E, Gabelle A, Thobois S*, et al.* C9orf72 repeat expansions are a rare genetic cause of parkinsonism. Brain 2013; 136: 385-391.

Levine TP, Daniels RD, Gatta AT, Wong LH and Hayes MJ. The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. Bioinformatics 2013; 29: 499-503.

Lin CL, Bristol LA, Jin L, Dykes-Hoberg M, Crawford T, Clawson L*, et al.* Aberrant RNA processing in a neurodegenerative disease: the cause for absent EAAT2, a glutamate transporter, in amyotrophic lateral sclerosis. Neuron 1998; 20: 589-602.

Lindquist SG, Duno M, Batbayli M, Puschmann A, Braendgaard H, Mardosiene S*, et al.* Corticobasal and ataxia syndromes widen the spectrum of C9ORF72 hexanucleotide expansion disease. Clin Genet 2013; 83: 279-283.

Liu EY, Russ J, Wu K, Neal D, Suh E, McNally AG*, et al.* C9orf72 hypermethylation protects against repeat expansion-associated pathology in ALS/FTD. Acta Neuropathol 2014;

Longstreth WT, McGuire V, Koepsell TD, Wang Y and van Belle G. Risk of amyotrophic lateral sclerosis and history of physical activity: a population-based case-control study. Arch Neurol 1998; 55: 201-206.

Ludolph AC, Hugon J, Dwivedi MP, Schaumburg HH and Spencer PS. Studies on the aetiology and pathogenesis of motor neuron diseases. 1. Lathyrism: clinical findings in established cases. Brain 1987; 110 (Pt 1): 149-165.

Mackenzie IRA, Bigio EH, Ince PG, Geser F, Neumann M, Cairns NJ*, et al.* Pathological TDP-43 distinguishes sporadic amyotrophic lateral sclerosis from amyotrophic lateral sclerosis with SOD1 mutations. Annals of Neurology 2007; 61: 427-434.

Mackenzie IRA, Rademakers R and Neumann M. TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. The Lancet Neurology 2010; 9: 995-1007.

Majounie E, Abramzon Y, Renton AE, Perry R, Bassett SS, Pletnikova O*, et al.* Repeat expansion in C9ORF72 in Alzheimer's disease. N Engl J Med 2012; 366: 283-284.

Majounie E, Renton AE, Mok K, Dopper EG, Waite A, Rollinson S*, et al.* Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. Lancet Neurol 2012; 11: 323-330.

Mann DM, Rollinson S, Robinson A, Bennion Callister J, Thompson JC, Snowden JS*, et al.* Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. Acta Neuropathol Commun 2013; 1: 68.

Mantovani A, Sica A, Sozzani S, Allavena P, Vecchi A and Locati M. The chemokine system in diverse forms of macrophage activation and polarization. Trends Immunol 2004; 25: 677-686.

May S, Hornburg D, Schludi MH, Arzberger T, Rentzsch K, Schwenk BM*, et al.* C9orf72 FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. Acta Neuropathol 2014;

McKee AC, Stein TD, Kiernan PT and Alvarez VE. The neuropathology of chronic traumatic encephalopathy. Brain Pathol 2015; 25: 350-364.

Millecamps S, Boillee S, Le Ber I, Seilhean D, Teyssou E, Giraudeau M*, et al.* Phenotype difference between ALS patients with expanded repeats in C9ORF72 and patients with mutations in other ALS-related genes. J Med Genet 2012; 49: 258-263.

Mizielinska S, Gronke S, Niccoli T, Ridler CE, Clayton EL, Devoy A*, et al.* C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. Science 2014;

Mizielinska S, Lashley T, Norona FE, Clayton EL, Ridler CE, Fratta P*, et al.* C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. Acta Neuropathol 2013; 126: 845-857.

Mok K, Traynor BJ, Schymick J, Tienari PJ, Laaksovirta H, Peuralinna T*, et al.* Chromosome 9 ALS and FTD locus is probably derived from a single founder. Neurobiol Aging 2012; 33: 209 e203-208.

Montuschi A, Iazzolino B, Calvo A, Moglia C, Lopiano L, Restagno G*, et al.* Cognitive correlates in amyotrophic lateral sclerosis: a population-based study in Italy. J Neurol Neurosurg Psychiatry 2014;

Mori K, Arzberger T, Grasser FA, Gijselinck I, May S, Rentzsch K*, et al.* Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. Acta Neuropathol 2013; 126: 881-893.

Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E*, et al.* The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science 2013; 339: 1335-1338.

Murray ME, Dejesus-Hernandez M, Rutherford NJ, Baker M, Duara R, Graff-Radford NR*, et al.* Clinical and neuropathologic heterogeneity of c9FTD/ALS associated with hexanucleotide repeat expansion in C9ORF72. Acta Neuropathol 2011; 122: 673-690.

Neumann M, Rademakers R, Roeber S, Baker M, Kretzschmar HA and Mackenzie IR. A new subtype of frontotemporal lobar degeneration with FUS pathology. Brain 2009; 132: 2922-2931.

Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT*, et al.* Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. Science 2006; 314: 130-133.

Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT*, et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science 2006; 314: 130-133.

Nicholson AM, Finch NA, Wojtas A, Baker MC, Perkerson RB, 3rd, Castanedes-Casey M*, et al.* TMEM106B p.T185S regulates TMEM106B protein levels: implications for frontotemporal dementia. J Neurochem 2013; 126: 781-791.

Nuytemans K, Bademci G, Kohli MM, Beecham GW, Wang L, Young JI*, et al.* C9ORF72 intermediate repeat copies are a significant risk factor for Parkinson disease. Ann Hum Genet 2013; 77: 351-363.

Pardo CA, Xu Z, Borchelt DR, Price DL, Sisodia SS and Cleveland DW. Superoxide dismutase is an abundant component in cell bodies, dendrites, and axons of motor neurons and in a subset of other neurons. Proc Natl Acad Sci U S A 1995; 92: 954-958.

Parton MJ, Broom W, Andersen PM, Al-Chalabi A, Nigel Leigh P, Powell JF*, et al.* D90A-SOD1 mediated amyotrophic lateral sclerosis: a single founder for all cases with evidence for a Cis-acting disease modifier in the recessive haplotype. Hum Mutat 2002; 20: 473.

Pesiridis GS, Lee VM-Y and Trojanowski JQ. Mutations in TDP-43 link glycine-rich domain functions to amyotrophic lateral sclerosis. Human Molecular Genetics 2009; 18: R156-R162.

Phukan J, Elamin M, Bede P, Jordan N, Gallagher L, Byrne S*, et al.* The syndrome of cognitive impairment in amyotrophic lateral sclerosis: a population-based study. J Neurol Neurosurg Psychiatry 2012; 83: 102-108.

Pickrell JK, Pai AA, Gilad Y and Pritchard JK. Noisy splicing drives mRNA isoform diversity in human cells. PLoS Genet 2010; 6: e1001236.

Pikkarainen M, Hartikainen P and Alafuzoff I. Neuropathologic features of frontotemporal lobar degeneration with ubiquitin-positive inclusions visualized with ubiquitin-binding protein p62 immunohistochemistry. J Neuropathol Exp Neurol 2008; 67: 280-298.

Premi E, Gazzina S, Bozzali M, Archetti S, Alberici A, Cercignani M, et al. Cognitive reserve in granulin-related frontotemporal dementia: from preclinical to clinical stages. PLoS One 2013; 8: e74762.

Preux PM, Couratier P, Boutros-Toni F, Salle JY, Tabaraud F, Bernet-Bernady P, et al. Survival prediction in sporadic amyotrophic lateral sclerosis. Age and clinical form at onset are independent risk factors. Neuroepidemiology 1996; 15: 153-160.

Puls I, Jonnakuty C, LaMonte BH, Holzbaur EL, Tokito M, Mann E, et al. Mutant dynactin in motor neuron disease. Nat Genet 2003; 33: 455-456.

Renton AE, Majounie E, Waite A, Simon-Sanchez J, Rollinson S, Gibbs JR*, et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 2011; 72: 257-268.

Renton AE, Majounie E, Waite A, Simon-Sanchez J, Rollinson S, Gibbs JR*, et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 2011; 72: 257-268.

Ritson GP, Custer SK, Freibaum BD, Guinto JB, Geffel D, Moore J*, et al.* TDP-43 mediates degeneration in a novel Drosophila model of disease caused by mutations in VCP/p97. J Neurosci 2010; 30: 7729-7739.

Roche JC, Rojas-Garcia R, Scott KM, Scotton W, Ellis CE, Burman R*, et al.* A proposed staging system for amyotrophic lateral sclerosis. Brain 2012; 135: 847-852.

Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A*, et al.* Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature 1993; 362: 59-62.

Rubino E, Rainero I, Chio A, Rogaeva E, Galimberti D, Fenoglio P*, et al.* SQSTM1 mutations in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Neurology 2012; 79: 1556-1562.

Sabatelli M, Conforti FL, Zollino M, Mora G, Monsurro MR, Volanti P*, et al.* C9ORF72 hexanucleotide repeat expansions in the Italian sporadic ALS population. Neurobiol Aging 2012; 33: 1848 e1815-1820.

Sareen D, O'Rourke JG, Meera P, Muhammad AK, Grant S, Simpkinson M*, et al.* Targeting RNA Foci in iPSC-Derived Motor Neurons from ALS Patients with a C9ORF72 Repeat Expansion. Sci Transl Med 2013; 5: 208ra149.

Sha SJ, Takada LT, Rankin KP, Yokoyama JS, Rutherford NJ, Fong JC*, et al.* Frontotemporal dementia due to C9ORF72 mutations: clinical and imaging features. Neurology 2012; 79: 1002-1011.

Simon-Sanchez J, Dopper EG, Cohn-Hokke PE, Hukema RK, Nicolaou N, Seelaar H*, et al.* The clinical and pathological phenotype of C9ORF72 hexanucleotide repeat expansions. Brain 2012; 135: 723-735.

Smith BN, Newhouse S, Shatunov A, Vance C, Topp S, Johnson L*, et al.* The C9ORF72 expansion mutation is a common cause of ALS+/-FTD in Europe and has a single founder. Eur J Hum Genet 2013; 21: 102-108.

Snowden JS, Rollinson S, Thompson JC, Harris JM, Stopford CL, Richardson AM*, et al.* Distinct clinical and pathological characteristics of frontotemporal dementia associated with C9ORF72 mutations. Brain 2012; 135: 693-708.

Spector DL and Lamond AI. Nuclear Speckles. Cold Spring Harbor Perspectives in Biology 2011; 3:

Spencer PS, Nunn PB, Hugon J, Ludolph AC, Ross SM, Roy DN*, et al.* Guam amyotrophic lateral sclerosis-parkinsonism-dementia linked to a plant excitant neurotoxin. Science 1987; 237: 517-522.

Sreedharan J, Blair IP, Tripathi VB, Hu X, Vance C, Rogelj B*, et al.* TDP-43 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis. Science 2008; 319: 1668-1672.

Stewart H, Rutherford NJ, Briemberg H, Krieger C, Cashman N, Fabros M*, et al.* Clinical and pathological features of amyotrophic lateral sclerosis caused by mutation in the C9ORF72 gene on chromosome 9p. Acta Neuropathol 2012; 123: 409-417.

Sutedja NA, Veldink JH, Fischer K, Kromhout H, Wokke JH, Huisman MH*, et al.* Lifetime occupation, education, smoking, and risk of ALS. Neurology 2007; 69: 1508-1514.

Thomas M, Alegre-Abarrategui J and Wade-Martins R. RNA dysfunction and aggrephagy at the centre of an amyotrophic lateral sclerosis/frontotemporal dementia disease continuum. Brain 2013; 136: 1345-1360.

Ticozzi N, Tiloca C, Calini D, Gagliardi S, Altieri A, Colombrita C*, et al.* C9orf72 repeat expansions are restricted to the ALS-FTD spectrum. Neurobiol Aging 2013;

Ticozzi N, Vance C, Leclerc AL, Keagle P, Glass JD, McKenna-Yasek D*, et al.* Mutational analysis reveals the FUS homolog TAF15 as a candidate gene for familial amyotrophic lateral sclerosis. Am J Med Genet B Neuropsychiatr Genet 2011; 156B: 285-290.

Traynor BJ, Codd MB, Corr B, Forde C, Frost E and Hardiman O. Incidence and prevalence of ALS in Ireland, 1995-1997: a population-based study. Neurology 1999; 52: 504-509.

Tresse E, Salomons FA, Vesa J, Bott LC, Kimonis V, Yao TP*, et al.* VCP/p97 is essential for maturation of ubiquitin-containing autophagosomes and this function is impaired by mutations that cause IBMPFD. Autophagy 2010; 6: 217-227.

Turner MR, Cagnin A, Turkheimer FE, Miller CC, Shaw CE, Brooks DJ*, et al.* Evidence of widespread cerebral microglial activation in amyotrophic lateral sclerosis: an [11C](R)-PK11195 positron emission tomography study. Neurobiol Dis 2004; 15: 601-609.

Tylleskar T, Banea M, Bikangi N, Cooke RD, Poulter NH and Rosling H. Cassava cyanogens and konzo, an upper motoneuron disease found in Africa. Lancet 1992; 339: 208-211.

van Blitterswijk M, Baker MC, DeJesus-Hernandez M, Ghidoni R, Benussi L, Finger E*, et al.* C9ORF72 repeat expansions in cases with previously identified pathogenic mutations. Neurology 2013; 81: 1332-1341.

van Blitterswijk M, DeJesus-Hernandez M, Niemantsverdriet E, Murray ME, Heckman MG, Diehl NN*, et al.* Association between repeat sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): a cross-sectional cohort study. Lancet Neurol 2013; 12: 978-988.

van Blitterswijk M, Mullen B, Heckman MG, Baker MC, DeJesus-Hernandez M, Brown PH*, et al.* Ataxin-2 as potential disease modifier in C9ORF72 expansion carriers. Neurobiol Aging 2014; 35: 2421 e2413-2427.

van Blitterswijk M, Mullen B, Nicholson AM, Bieniek KF, Heckman MG, Baker MC*, et al.* TMEM106B protects C9ORF72 expansion carriers against frontotemporal dementia. Acta Neuropathol 2014;

van Blitterswijk M, van Es MA, Hennekam EA, Dooijes D, van Rheenen W, Medic J, et al. Evidence for an oligogenic basis of amyotrophic lateral sclerosis. Hum Mol Genet 2012; 21: 3776-3784.

Van Deerlin VM, Sleiman PM, Martinez-Lage M, Chen-Plotkin A, Wang LS, Graff-Radford NR, et al. Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. Nat Genet 2010; 42: 234-239.

van der Zee J, Gijselinck I, Dillen L, Van Langenhove T, Theuns J, Engelborghs S*, et al.* A Pan-European Study of the C9orf72 Repeat Associated with FTLD: Geographic Prevalence, Genomic Instability, and Intermediate Repeats. Human Mutation 2013; 34: 363-373.

van Rheenen W, van Blitterswijk M, Huisman MH, Vlam L, van Doormaal PT, Seelen M*, et al.* Hexanucleotide repeat expansions in C9ORF72 in the spectrum of motor neuron diseases. Neurology 2012; 79: 878-882.

Vance C, Rogelj B, Hortobagyi T, De Vos KJ, Nishimura AL, Sreedharan J*, et al.* Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science 2009; 323: 1208-1211.

Veldink JH, Kalmijn S, Groeneveld GJ, Titulaer MJ, Wokke JH and van den Berg LH. Physical activity and the association with sporadic ALS. Neurology 2005; 64: 241-245.

Waite AJ, Baumer D, East S, Neal J, Morris HR, Ansorge O*, et al.* Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. Neurobiol Aging 2014; 35: 1779 e1775-1779 e1713.

Wicks P, Abrahams S, Papps B, Al-Chalabi A, Shaw CE, Leigh PN*, et al.* SOD1 and cognitive dysfunction in familial amyotrophic lateral sclerosis. Journal of Neurology 2009; 256: 234-241.

Wu CH, Fallini C, Ticozzi N, Keagle PJ, Sapp PC, Piotrowska K*, et al.* Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. Nature 2012; 488: 499-503.

Xi Z, Rainero I, Rubino E, Pinessi L, Bruni AC, Maletta RG*, et al.* Hypermethylation of the CpG-island near the C9orf72 G4C2-repeat expansion in FTLD patients. Hum Mol Genet 2014;

Xi Z, Yunusova Y, van Blitterswijk M, Dib S, Ghani M, Moreno D*, et al.* Identical twins with the C9orf72 repeat expansion are discordant for ALS. Neurology 2014; 83: 1476-1478.

Xi Z, Zinman L, Moreno D, Schymick J, Liang Y, Sato C*, et al.* Hypermethylation of the CpG Island Near the G4C2 Repeat in ALS with a C9orf72 Expansion. Am J Hum Genet 2013; 92: 981-989.

Zhao W, Beers DR, Henkel JS, Zhang W, Urushitani M, Julien JP*, et al.* Extracellular mutant SOD1 induces microglial-mediated motoneuron injury. Glia 2010; 58: 231-243.

Zu T, Gibbens B, Doty NS, Gomes-Pereira M, Huguet A, Stone MD*, et al.* Non-ATG-initiated translation directed by microsatellite expansions. Proc Natl Acad Sci U S A 2011; 108: 260-265.

Zu T, Liu Y, Banez-Coronel M, Reid T, Pletnikova O, Lewis J*, et al.* RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. Proc Natl Acad Sci U S A 2013;

7. Full list of published work

h-index: **10**
i10-index: **10**

1. **Cooper-Knock J***, Kirby J*, Highley JR and Shaw PJ. The spectrum of C9orf72 mediated neurodegeneration and amyotrophic lateral sclerosis (ALS). **Neurotherapeutics**. 2015 Mar 3.
2. **Cooper-Knock J**, Bury JJ, Heath PR, Wyles M, Higginbottom A, Walsh MJ, Gelsthorpe C, Highley JR, Hautbergue GM, Rattray M, Kirby J, Shaw PJ. *C9ORF72* GGGGCC expanded repeats produce splicing dysregulation which correlates with disease severity in amyotrophic lateral sclerosis. **PLoS One** [under review]
3. **Cooper-Knock J**, Higginbottom A, Stopford MJ, Highley JR, Ince PG, Wharton SB, Kirby J, Hautbergue GM and Shaw PJ. Antisense RNA foci in the motor neurons of C9ORF72-ALS patients are associated with TDP-43 proteinopathy **Acta Neuropathologica** [under review]
4. Beer AM, **Cooper-Knock J**, Higginbottom A, Highley JR, Wharton SB, Ince PG, Milano A, Jones AA, Al-Chalabi A, Kirby J, Shaw PJ. Intermediate length C9orf72 expansion in an ALS patient without classical C9orf72 neuropathology. **Amyotroph Lateral Scler Frontotemporal Degener.** 2014 Dec 1:1-3
5. Bayatti N, **Cooper-Knock J**, Bury JJ, Wyles M, Heath PR, Kirby J, Shaw PJ. Comparison of blood RNA extraction methods used for gene expression profiling in amyotrophic lateral sclerosis. **PLoS One**. 2014;9:e87508.
6. **Cooper-Knock J**, Shaw PJ, Kirby J. The widening spectrum of C9ORF72-related disease; genotype/phenotype correlations and potential modifiers of clinical phenotype. **Acta Neuropathol.** 2014;127:333-45.
7. Green NH, Nicholls Z, Heath PR, **Cooper-Knock J**, Corfe BM, MacNeil S, Bury JP. Pulsatile exposure to simulated reflux leads to changes in gene expression in a 3D model of oesophageal mucosa. **International Journal of Experimental Pathology**. 2014 95:216-28

8. Highley JR, Kirby J, Jansweijer JA, Webb PS, Hewamadduma CA, Heath PR, Higginbottom A, Raman R, Ferraiuolo L, **Cooper-Knock J**, McDermott CJ, Wharton SB, Shaw PJ, Ince PG. Loss of nuclear TDP-43 in ALS causes altered expression of splicing machinery and widespread dysregulation of RNA splicing in motor neurons. **Neuropathology and Applied Neurobiology**. 2014 40:670-85

9. **Cooper-Knock J+,** Walsh MJ+, Higginbottom A, Highley JR, Dickman MJ, Edbauer D, Ince PG, Wharton SB, Wilson SA, Kirby J, Hautbergue GM, Shaw PJ. Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. **Brain**. 2014; 137:2040-51

10. Walsh MJ+, **Cooper-Knock J+,** Dodd JE, Stopford MJ, Mihaylov SR, Kirby J, Shaw PJ, Hautbergue GM. Decoding the pathophysiological mechanisms that underlie RNA dysregulation in neurodegenerative disorders: a review of the current state of the art. **Neuropathology and Applied Neurobiology**. 2014 Epub 1st October

11. Gallagher MD, Suh E, Grossman M, Elman L, McCluskey L, Van Swieten JC, Al-Sarraj S, Neumann M, Gelpi E, Ghetti B, Rohrer JD, Halliday G, Van Broeckhoven C, Seilhean D, Shaw PJ, Frosch MP, Alafuzoff I, Antonell A, Bogdanovic N, Brooks W, Cairns NJ, **Cooper-Knock J**, Cotman C, Cras P, Cruts M, De Deyn PP, Decarli C, Dobson-Stone C, Engelborghs S, Fox N, Galasko D, Gearing M, Gijselinck I, Grafman J, Hartikainen P, Hatanpaa KJ, Highley JR, Hodges J, Hulette C, Ince PG, Jin LW, Kirby J, Kofler J, Kril J, Kwok JB, Levey A, Lieberman A, Llado A, Martin JJ, Masliah E, McDermott CJ, McKee A, McLean C, Mead S, Miller CA, Miller J, Munoz DG, Murrell J, Paulson H, Piguet O, Rossor M, Sanchez-Valle R, Sano M, Schneider J, Silbert LC, Spina S, van der Zee J, Van Langenhove T, Warren J, Wharton SB, White Iii CL, Woltjer RL, Trojanowski JQ, Lee VM, Van Deerlin V, Chen-Plotkin AS. TMEM106B is a genetic modifier of frontotemporal lobar degeneration with C9orf72 hexanucleotide repeat expansions. **Acta Neuropathol**. 2014;127:407-18.

12. Peters OM, Shelkovnikova T, Highley JR, **Cooper-Knock J**, Hortobágyi T, Troakes C, Natalia Ninkina N, Buchman VL. Gamma-synuclein pathology in

amyotrophic lateral sclerosis. **Annals of Clinical and Translational Neurology.** Epub 1st November 2014

13. A Ismail+, **J Cooper-Knock+**, JR Highley, A Milano, J Kirby , J Lowe, CS Constantinescu, SJ Walters, S Price, CJ McDermott, S Sawcer, DAS Compston, B Sharrack, PJ Shaw. Concurrence of multiple sclerosis and amyotrophic lateral sclerosis in patients with hexanucleotide repeat expansions of C9ORF72. **J Neurol Neurosurg Psychiatry**. 2013 Jan;84:79-87

14. Lo C, **Cooper-Knock J**, Garrard K, Martindale J, Williams T, Shaw P. Concurrent amyotrophic lateral sclerosis and cystic fibrosis supports common pathways of pathogenesis. Amyotroph Lateral Scler Frontotemporal Degener. 2013 Jan 4. **Amyotroph Lateral Scler Frontotemporal Degener.** 14:473-475

15. Jones AR, Woollacott I, Shatunov A, **Cooper-Knock J**, Buchman V, Sproviero W, Smith B, Scott KM, Balendra R, Abel O, McGuffin P, Ellis CM, Shaw PJ, Morrison KE, Farmer A, Lewis CM, Leigh PN, Shaw CE, Powell JF, Al-Chalabi A. Residual association at C9orf72 suggests an alternative ALS-causing hexanucleotide repeat. **Neurobiology of Aging** 34:2234

16. Buchman VL, **Cooper-Knock** J, Connor-Robson N, Higginbottom A, Kirby J, Razinskaya OD, Ninkina N, Shaw PJ. Simultaneous and independent detection of C9ORF72 alleles with low and high number of GGGGCC repeats using an optimised protocol of Southern blot hybridisation. **Molecular Neurodegeneration**. 2013; 8:1-6

17. **Cooper-Knock, J.**, Frolov A., HighleyJ.R., Charlesworth G., Kirby J., Milano A.*, et al*. *C9ORF72* expansions, parkinsonism, and Parkinson disease: a clinicopathologic study. **Neurology** 2013; 81: 808-811.

18. **Cooper-Knock, J.**, A. Higginbottom, N. Connor-Robson, N. Bayatti, J. J. Bury, J. Kirby*, et al*. *C9ORF72* transcription in a frontotemporal dementia case with two expanded alleles. **Neurology** 2013; 81:1719-21

19. **Cooper-Knock J**+, Hewitt C+, Highley JR, Brockington A, Milano A, Man S, Martindale J, Hartley J , Walsh T, Gelsthorpe C, Baxter L, Forster G, Fox M, Mok K, McDermott CJ, Traynor B, Kirby J, Hardy J, Wharton SB, Ince PG,

Shaw PJ. Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. **Brain**. 2012;135:751-64.

20. Majounie E, Renton AE, Mok K, Nicalou N, Waite A, Rollinson S, Chiò A, Restagno G, Simon-Sanchez J, van Swieten J, Abramzon Y, Johnson JO, Sendtner M, Pamphlett R,. Orrell RW, Mead S, Houlden H, Rohrer JD, Morrison K, Talbot K, Ansorge O, The Chromosome 9-ALS/FTD Consortium (including **Cooper-Knock J**), The ITALSGEN Consortium, Englund E, Borghero G, McCluskey L, Trojanowski JQ, van Deerlin VM, Schellenberg GD, Nalls GA, Drory V, Brice A, Drepper C , Williams N, Kirby J, Shaw P, Hardy J, Singleton A, Tienari PJ, Heutink P, Morris H, Pickering-Brown A, Traynor BJ Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. **Lancet Neurol.** 2012;11:323-30.
21. Goodall E, Bury J, **Cooper-Knock J**, Shaw PJ, Kirby J. Genetics of Familial Amyotrophic Lateral Sclerosis. Amyotrophic Lateral Sclerosis, ISBN 979-953-307-199-1 **InTech** 2011
22. **Cooper-Knock J**, Jenkins T, Shaw PJ. Clinical and Molecular Aspects of Motor Neuron Disease. Colloquium Series on Genomic and Molecular Medicine. ISBN: 9781615044283 **Morgan & Claypool Life Sciences** 2013.
23. **Cooper-Knock J**, Kirby J, Heath P, Shaw PJ. Gene Expression Profiling in Discovery of Pathophysiology in Neurodegenerative Disease: A Review **Nature Reviews Neurology** 2012; 8:518-30
24. **Cooper-Knock J,** Bury J, Ferraiuolo L, Goodall E, Shaw PJ, Kirby J. Insights Arising from Gene Expression Profiling in Amyotrophic Lateral Sclerosis. Amyotrophic Lateral Sclerosis, ISBN 979-953-307-199-1 **InTech** 2011
25. **Cooper-Knock J**, Pepper I, Hodgson T, Sharrack B. Early diagnosis of Horner's syndrome using topical apraclonidine. **Journal of Neuro-Ophthalmology** 2011:214-6.
26. **Cooper-Knock J**, Ahmedzai SH, Shaw PJ. The use of subcutaneous glycopyrrolate in management of sialorrhea in bulbar motor neuron disease and facilitating the use of non-invasive ventilation **Amyotrophic Lateral Sclerosis** 2011 12:464-5.

27. **Cooper-Knock J**, Wood L, Pearson Z, Tate M, Shaw PJ. Neurology and Clinical Neuroanatomy: The Basics **Hodder Arnold** 2014

28. Hu M, Butterworth R, Kumar V, **Cooper J ***, Jones E, Catterall L, Ben-Shlomo Y. How common and what are the determinants of sub-optimal care for Parkinson's disease patients: the Milton Keynes community study. **Parkinsonism and Related Disorders** 2011 17:177-81

29. Hu M, Cooper J *, Beamish R, Jones E, Butterworth R, Catterall L, Ben-Shlomo Y. How well do we recognise non-motor symptoms in a British Parkinson's disease population? **J Neurol** 2011 258:1513-7